Technology Survival Manual

Index

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.*io*

Technology Survival Manual

Introduction

One of the difficulties of getting to grips with Planet Limited vision is that though it is very simple, it is based on a vast range of underlying technologies and specialised practices. Technology practitioners, like ourselves, tend to forget that a lot of what we take for granted can be completely new information to the people we talk to and work with.

This brief is intended to provide the reader with an understanding of the various technologies, practices and standards that make or made our work.

We really need to thank the Wikipedia.org community for making this document possible.

Types Of Information

We provide information in this document as per the following broad classifications;

- General Design Considerations
- Mobile Telephony
- E-Commerce
- The Internet
- Data & Networking Practises
- Other

We explore each of the topics extensively, albeit only on sub-topics that have a direct impact on the development of SMS-B. The sub-topics we covers are as follows;

- **General Design Considerations:** Back End Systems, Content & Mobile Devices

- **Mobile Telephony:** SMS, MMS, Short Codes, Mobile Marketing, Mobile Learning, Mobile News, Mobile Publishing, Mobile Blogs, GPS2SMS, Reverse SMS Billing, SMS Gateways, Short Message Service Centre, SMPP, Mobile Phone, IMEI, SIM and GSM.

- **E-Commerce:** Electronic Commerce, Online Banking, Electronic Funds Transfer (EFT), Electronic Money, Mobile Commerce, Mobile Banking, Mobile Payments and Smartcards etc

- **The Internet:** The Internet, The World Wide Web, URLs, Domain Names, The Domain Name System, Websites, Meta Elements, Search Engines, E-Mail and Instant Messaging etc

- **Data & Networking Practises:** Computer Networks, Network Protocol, Ambient Network, Replication, File Sharing, Distributed Computing, Grid Computing, P2P and Servents etc

- **Other:** Social Networking, Six Degrees of Separation, SMS Language & Abbreviations and Push & Pull Marketing Strategy etc.

Design Considerations – Back-End Systems

Overview

What is critical to the success of our concepts is the availability, quality, reliability and efficiency of the processing of information, data and media submitted to and from the end users via the input points.

To achieve them we will borrow concepts and methodologies from the field of supercomputing and employ them in a manner that could arguably be described as logical poetry.

In this section we detail some of the critical issues relating to the problem the solution must resolve.

Availability

Mobile based business models are dependent on the availability and speed of response to all parties in the transaction.

In computer science, the term availability has the following meanings:

1. The degree to which a system, subsystem, or equipment is operable and in a committable state at the start of a mission, when the mission is called for at an unknown, i.e., a random, time. Simply put, availability is the proportion of time a system is in a functioning condition.

    - Note 1: The conditions determining operability and committability must be specified.

    - Note 2: Expressed mathematically, availability is 1 minus the unavailability.

2. The ratio of (a) the total time a functional unit is capable of being used during a given interval to (b) the length of the interval.

    - Note 1: An example of availability is 100/168 if the unit is capable of being used for 100 hours in a week.

    - Note 2: Typical availability objectives are specified either in decimal fractions, such as 0.9998, or sometimes in a logarithmic unit called nines, which corresponds roughly to a number of nines following the decimal point, such as "five nines" for 0.99999 reliability.

Source: from Federal Standard 1037C in support of MIL-STD-188

Please see the appendices of this document for more information about this topic.

Scalability

The customer's ventures are potentially massive in scale. In order to remain both sustainable and profitable, take must be taken to enable the system to scale both effectively and efficiently.

Scalability is a desirable property of a system, a network, or a process, which indicates its ability to either handle growing amounts of work in a graceful manner, or to be readily enlarged. For example, it can refer to the capability of a system to increase total throughput under an increased load when resources (typically hardware) are added. An analogous meaning is implied when the word is used in a commercial context, where scalability of a company implies that the underlying business model offers the potential for economic growth within the company.

Scalability, as a property of systems, is generally difficult to define and in any particular case it is necessary to define the specific requirements for scalability on those dimensions which are deemed important. It is a highly significant issue in electronics systems, database, routers, and networking. A system, whose performance improves after adding hardware, proportionally to the capacity added, is said to be a scalable system. An algorithm, design, networking protocol, program, or other system is said to scale if it is suitably efficient and practical when applied to large situations (e.g. a large input data set or large number of participating nodes in the case of a distributed system). If the design fails when the quantity increases then it does not scale.

Dimensions

Scalability can be measured in various dimensions, such as:

- Load scalability: The ability for a distributed system to easily expand and contract its resource pool to accommodate heavier or lighter loads. Alternatively, the ease with which a system or component can be modified, added, or removed, to accommodate changing load.
- Geographic scalability: The ability to maintain performance, usefulness, or usability regardless of expansion from concentration in a local area to a more distributed geographic pattern.
- Administrative scalability: The ability for an increasing number of organizations to easily share a single distributed system.

Add As Needed

It is likely that the growth of the customer’s requirements will be opportunity based – meaning that linear or predictable system expansion planning will at best be costing guidelines and, at worst, irrelevant.

As the customer’s current scope of operation is potentially wide and diverse both related to services and geographical location, the system is required to be readied for incremental, unplanned and organic growth from its infancy. This requires the solution to be able to be strengthened as required on any point of its logical model requiring additional capacity on a manageable, as needed basis.

This is distinguished from the typical scaling requirements in the fact that the scaling of capacity upwards cannot generally occur system wide in a parallel, holistic fashion.

Rapid Scaling

The time frame in which the customer operates their business is measured in days. This in itself presents issues around the speed at which capacities can be improved and presents few luxuries to system planners and architects.

The system must be able to be scaled both rapidly and on a “as needed” basis, while not disrupting any of its current, core or existing functions at the time.

In short, the ideal system’s rapidity of scale should be related to the availability of hardware and the logistics surrounding its implementation. The software portion of the system’s availability requires readiness in days.

Cost Efficiency

Cost is a major driver of any business model, plan or project but in the context of interactive mobile based businesses it presents a unique opportunity.

Efficiency is directly related to the volume of transactions processed by the business, however increased processing has an associated effort and equipment cost. Even though technology alleviates the effort and equipment prices are constantly improving directly in proportion to the capacity provided, take must still be taken so that massive or unpredictable growth does not require an equally massive outlay of capital while at the same time not disrupting existing operations (which have a direct, detrimental impact on earnings).

Self-Monitoring

Hardware capacity is finite and as growth of the processing volumes is highly likely, the system is required to be able to identify performance lags that can be addressed by additional hardware.

The system requires a reliable component which can monitor the performance of the system. This is necessary in as far as if volumes increase by whatever ratio over time, the system is able to identify it requirements for additional processing, memory and bandwidth.

This module is required to distinguish between performance measurements related to information received from the GSM input points and the system itself.

Strategies For Resolving Issues

In order to solve all of the requirements we are going to borrow methodologies from the field of supercomputing and integrate elements of them into the Foundation System. The specific methodologies we are going to employ are as follows:

- High Availability Protocols
- Distributed Computing
- Parallel Computing
- N-Tier Architecture
- Computer Clusters
- Grid Computing
- Data Grids
- Grid File System
- MIMD

We will very briefly review each of these methodologies and explain what they mean.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited*****.io***

High Availability Protocol

High availability is a system design protocol and associated implementation that ensures a certain absolute degree of operational continuity during a given measurement period.

Availability refers to the ability of the user community to access the system, whether to submit new work, update or alter existing work, or collect the results of previous work. If a user cannot access the system, it is said to be unavailable. Generally, the term downtime is used to refer to periods when a system is unavailable.

Planned & Unplanned

A distinction needs to be made between planned downtime and unplanned downtime. Typically, planned downtime is a result of maintenance that is disruptive to system operation and usually cannot be avoided with a currently installed system design. Planned downtime events might include patches to system software that require a reboot or system configuration changes that only take effect upon a reboot. In general, planned downtime is usually the result of some logical, management-initiated event. Unplanned downtime events typically arise from some physical event, such as a hardware failure or environmental anomaly. Examples of unplanned downtime events include power outages, failed CPU or RAM components (or possibly other failed hardware components), an over-temperature related shutdown, logically or physically severed network connections, catastrophic security breaches, or various application, middleware, and operating system failures.

Many computing sites exclude planned downtime from availability calculations, assuming, correctly or incorrectly, that planned downtime has little or no impact upon the computing user community. By excluding planned downtime, many systems can claim to have phenomenally high availability, which might give the illusion of continuous availability. Systems that exhibit truly continuous availability are comparatively rare and higher priced, and they have carefully implemented specialty designs that eliminate any single point of failure and allow online hardware, network, operating system, middleware, and application upgrades, patches, and replacements.

Percentage Calculation

Availability is usually expressed as a percentage of uptime in a given year. In a given year, the number of minutes of unplanned downtime are tallied for a system; the aggregate unplanned downtime is divided by the total number of minutes in a year (approximately 525,600), producing a percentage of downtime; the complement is the percentage of uptime, which is what is typically referred to as the availability of the system. Common values of availability, typically stated as a number of "nines", for highly available systems are:

- 99.9% ≡ 43.8 minutes/month or 8.76 hours/year ("three nines")
- 99.99% ≡ 4.38 minutes/month or 52.6 minutes/year ("four nines")
- 99.999% ≡ 0.44 minutes/month or 5.26 minutes/year ("five nines")

It should be noted that uptime and availability are not synonymous. A system can be up, but not available, as in the case of a network outage.

System Design For High Availability

Paradoxically, adding more components to an overall system design can actually undermine efforts to achieve high availability. That's because complex systems inherently have more

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

potential failure points and are more difficult to implement correctly. The most highly available systems hew to a simple design pattern: a single, high quality, multi-purpose physical system with comprehensive internal redundancy running all interdependent functions paired with a second, like system at a separate physical location. This classic design pattern is common among financial institutions, for example.

Distributed Computing

Distributed Computing is a method of computer processing in which different parts of a program run simultaneously on two or more computers that are communicating with each other over a network. Distributed computing is a type of segmented or parallel computing. But the latter term is most commonly used to refer to processing in which different parts of a program run simultaneously on two or more processors that are part of the same computer. While both types of processing require that a program be segmented—divided into sections that can run simultaneously, distributed computing also requires that the division of the program take into account the different environments on which the different sections of the program will be running. For example, two computers are likely to have different file systems and different hardware components.

Parallel Computing

Parallel Computing is the simultaneous execution of some combination of multiple instances of programmed instructions and data on multiple processors in order to obtain results faster. The idea is based on the fact that the process of solving a problem usually can be divided into smaller tasks, which may be carried out simultaneously with some coordination.

N-Tier Architecture

N-Tier Architecture is a client-server architecture in which an application is executed by more than one distinct software agent. For example, an application that uses middleware to service data requests between a user and a database employs multi-tier architecture. The most widespread use of "multi-tier architecture" refers to three-tier architecture

Computer Clusters

A Computer Cluster is a group of tightly coupled computers that work together closely so that in many respects they can be viewed as though they are a single computer. The components of a cluster are commonly, but not always, connected to each other through fast local area networks. Clusters are usually deployed to improve performance and/or availability over that provided by a single computer, while typically being much more cost-effective than single computers of comparable speed or availability.

Grid Computing

Grid Computing is a phrase in distributed computing which can have several meanings:

- A local computer cluster which is like a "grid" because it is composed of multiple nodes.
- Offering online computation or storage as a metered commercial service, known as utility computing, "computing on demand", or "cloud computing".
- The creation of a "virtual supercomputer" by using spare computing resources within an organization.

- The creation of a "virtual supercomputer" by using a network of geographically dispersed computers. Volunteer computing, which generally focuses on scientific, mathematical, and academic problems, is the most common application of this technology.

These varying definitions cover the spectrum of "distributed computing", and sometimes the two terms are used as synonyms. Functionally, one can also speak of several types of grids:

- Computational grids (including CPU Scavenging grids) which focuses primarily on computationally-intensive operations.
- Data grids or the controlled sharing and management of large amounts of distributed data.
- Equipment grids which have a primary piece of equipment e.g. a telescope, and where the surrounding Grid is used to control the equipment remotely and to analyze the data produced.

Data Grids

A Data Grid is a grid computing system that deals with data — the controlled sharing and management of large amounts of distributed data. These are often, but not always, combined with computational grid computing systems.

Grid File Systems

A Grid File System is a computer file system whose goal is improved reliability and availability by taking advantage of many smaller file storage areas.

MIMD

MIMD (Multiple Instruction stream, Multiple Data stream) is a technique employed to achieve parallelism. Machines using MIMD have a number of processors that function asynchronously and independently. At any time, different processors may be executing different instructions on different pieces of data. MIMD architectures may be used in a number of application areas such as computer-aided design/computer-aided manufacturing, simulation, modelling, and as communication switches.

MIMD machines can be of either shared memory or distributed memory categories. These classifications are based on how MIMD processors access memory. Shared memory machines may be of the bus-based, extended, or hierarchical type. Distributed memory machines may have hypercube or mesh interconnection schemes.

Information Quality

Information quality (IQ) is a term to describe the quality of the content of information systems. Most information system practitioners use the term synonymously with data quality. However, as many academics make a distinction between data and information, some will insist on a distinction between data quality and information quality. Information quality assurance is confidence that particular information meets some context specific quality requirements.

"Information quality" is a measure of the value which the information provides to the user of that information. 'Quality' is subjective and the quality of information can vary among users and among uses of the information. Furthermore, accuracy is just one element of IQ and this can be source-dependent. Often there is a trade-off between accuracy and other aspects of the information determining its suitability for any given tasks.

Dimensions of Information Quality

The generally accepted list of elements used in assessing subjective Information Quality are those put forth in Wang & Strong (1996).:

- Intrinsic IQ: Accuracy, Objectivity, Believability, Reputation
- Contextual IQ: Relevancy, Value-Added, Timeliness, Completeness, Amount of information
- Representational IQ: Interpretability, Ease of understanding, Concise representation, Consistent representation
- Accessibility IQ: Accessibility, Access security

Data Sets

A data set (or dataset) is a collection of data, usually presented in tabular form. Each column represents a particular variable, and each row is an assignment of values for each of the variables to a member of the set in question.

In the simplest case, there is only one variable, and then the data set consists of a single column of values, often represented as a list.

The values may be numbers, such as real numbers or integers, for example representing a person's height in centimetres, but may also be nominal data (i.e., not consisting of numerical values), for example representing a person's ethnicity. For each variable, the values will normally all be of the same kind. However, there may also be "missing values", which need to be indicated in some way.

In statistics data sets usually come from actual observations obtained by sampling a statistical population, and each row corresponds to the observations on one element of that population. Data sets may further be generated by algorithms for the purpose of testing certain kinds of software.

While the term suggests a relationship to set theory it should not be assumed that a given data set is, in fact, a set in the usual mathematically sense. The rows of a data set need not be distinct, and so a data set is technically a multiset.

Data Quality

Data Quality refers to the quality of data. Data are of high quality "if they are fit for their intended uses in operations, decision making and planning" (J.M. Juran). Alternatively, the data are deemed of high quality if they correctly represent the real-world construct to which they refer. These two views can often be in disagreement, even about the same set of data used for the same purpose.

Some definitions of Data Quality include;

- Data Quality refers to the degree of excellence exhibited by the data in relation to the portrayal of the actual phenomena. GIS Glossary
- The state of completeness, validity, consistency, timeliness and accuracy that makes data appropriate for a specific use. Government of British Columbia

- The totality of features and characteristics of data that bears on their ability to satisfy a given purpose; the sum of the degrees of excellence for factors related to data.

Overview Of Data Quality

There are a number of theoretical frameworks for understanding data quality. One framework seeks to integrate the product perspective (conformance to specifications) and the service perspective (meeting consumers' expectations) (Kahn et al 2002). Another framework is based in semiotics to evaluate the quality of the form, meaning and use of the data (Price and Shanks, 2004). One highly theoretical approach analyzes the ontological nature of information systems to define data quality rigorously (Wand and Wang, 1996).

A considerable amount of data quality research involves investigating and describing various categories of desirable attributes (or dimensions) of data. These lists commonly include accuracy, correctness, currency, completeness and relevance. Nearly 200 such terms have been identified and there is little agreement in their nature (are these concepts, goals or criteria?), their definitions or measures (Wang et al, 1993). Software engineers may recognize this as a similar problem to the so-called Ilities.

MIT has a Total Data Quality Management program, led by Professor Richard Wang, which produces a large number of publications and hosts a significant international conference in this field.

In practice, data quality is a concern for professionals involved with a wide range of information systems, ranging from data-warehousing and business intelligence to customer relationship management and supply chain management. One industry study estimated the total cost to the US economy of data quality problems at over US$600 billion per annum (Eckerson, 2002). In fact, the problem is such a concern that companies are beginning to set up a data governance team whose sole role in the corporation is to be responsible for data quality. In some organizations, this data governance function has been established as part of a larger Regulatory Compliance function - a recognition of the importance of Data/Information Quality to organizations

Problems with data quality don't only arise from incorrect data. Inconsistent data is a problem as well. Eliminating data shadow systems and centralizing data in a warehouse is one of the initiatives a company can take to ensure data consistency.

The market is going some way to providing data quality assurance. A number of vendors make tools for analyzing and repairing poor quality data in situ, service providers can clean the data on a contract basis and consultants can advise on fixing processes or systems to avoid data quality problems in the first place. Most data quality tools offer a series of tools for improving data, which may include some or all of the following:

- Data profiling - initially assessing the data to understand its quality challenges
- Data standardization - a business rules engine that ensures that data conforms to quality rules
- Geocoding - for name and address data. Corrects data to US and Worldwide postal standards
- Matching or Linking - a way to compare data so that similar, but slightly different records can be aligned. Matching may use "fuzzy logic" to find duplicates in the data. It often recognizes that 'Bob' and 'Robert' may be the same individual. It might be able to manage 'householding', or finding links between husband and wife at the same address, for example. Finally, it often can build a 'best of breed' record, taking

the best components from multiple data sources and building a single super-record.
- Monitoring - keeping track of data quality over time and reporting variations in the quality of data.
- Batch and Real time - Once the data is initially cleansed (batch), companies often want to build the processes into enterprise applications to keep it clean.

There are several well-known authors and self-styled experts, with Larry English perhaps the most popular guru. In addition, the International Association for Information and Data Quality (IAIDQ) was established in 2004 to provide a focal point for professionals and researchers in this field.

Data Cleansing

Data cleansing is the act of detecting and correcting (or removing) corrupt or inaccurate records from a record set.

After cleansing, a data set will be consistent with other similar data sets in the system. The inconsistencies detected or removed may have been originally caused by different data dictionary definitions of similar entities in different stores, may have been caused by user entry errors, or may have been corrupted in transmission or storage.

Pre-processing the data will also guarantee that it is unambiguous, correct, and complete.

The actual process of data cleansing may involve removing typos or validating and correcting values against a known list of entities. The validation may be strict (such as rejecting any address that does not have a valid ZIP code) or fuzzy (such as correcting records that partially match existing, known records).

Data cleansing is synonymous with the less frequently-used term data scrubbing. Data cleansing differs from data validation in that validation almost invariably means data is rejected from the system at entry and is performed at entry time, rather than on batches of data.

Missing Values

In statistics, missing values are a common occurrence. Several statistical methods have been developed to deal with this problem. Missing values mean that no data value is stored for the variable in the current observation. Modern statistical packages made dealing with missing values much easier. Because a large volume of critical data will be feeding the solution, it is important that we take into account methods to deal with incomplete and even missing data sets.

Some techniques of dealing with missing values include

- Imputation (statistics)
    - Single imputation
    - Multiple imputation
    - EM imputation (also known as expectation-maximization imputation, see Expectation-maximization algorithm)
- Full information maximum likelihood estimation
- Indicator variable
- List wise deletion/case wise deletion
- Pair wise deletion

- Mean substitution
- Mplus
- MCAR (missing completely at random)

Imputation

Imputation is the substitution of some value for a missing data point or a missing component of a data point. Once all missing values have been imputed, the dataset can then be analysed using standard techniques for complete data. The analysis should ideally take into account that there is a greater degree of uncertainty than if the imputed values had actually been observed, however, and this generally requires some modification of the standard complete-data analysis methods. While many imputation techniques are available, two of the most commonly used are hot-deck imputation and regression imputation.

Hot-deck imputation fills in missing values on incomplete records using values from similar, but complete records of the same dataset. (The term "hot deck" dates back to the storage of data on punch cards, and indicates that the information donors come from the same dataset as the recipients; the stack of cards was hot because it was currently being processed. Cold-deck imputation, by contrast, selects donors from another dataset.)

Since standard analysis techniques do not reflect the additional uncertainty due to imputing for missing data, further adjustments (such as multiple imputation or a Rao-Shao correction) are necessary to account for this.

Imputation is not the only method available for handling missing data. It usually gives better results than list wise deletion (in which all subjects with any missing values are omitted from the analysis), and may be competitive with a maximum likelihood approach in many circumstances.

Data Maintenance

Data maintenance is the adding, deleting, changing and updating of binary and high level files, and the real world data associated with those files. Data can be maintained manually and/or through an automated program, but at origination and translation/delivery point must be translated into a binary representation for storage. Data is usually edited at a slightly higher level in a format relevant to the content of the data (such as text, images, or scientific or financial information).

Design Considerations – Content

High-Level

In this section we concern ourselves with the various concepts and criteria relating to the management and distribution of content, be it media, music, video, games etc. Specifically we look towards the various disciplines in the practise and illuminate these for future consideration as the central control centre design is clarified.

Why Bother?

We have to take care that those content vendors that we engage can demonstrably establish our competence and compliance in accordance with their particular regime of Digital Rights Management. As these are likely to have a royalty and billing impact on the company the value hereof should not be understated. Further, a sound basis of practise will ensure that the user experience, specifically in regards to user generated content, is both predictable and reliable.

This sound foundation will not only allow for useful business statistics in respect of capacity and delivery planning (internal) but also form a basis for user/media interaction statistics (external) which will hold great value to content vendors. The information gathered from both will allow us to continually improve both our service and media offered to our customers via the mobile browser.

To this end we will review the relevant concepts and practises in the balance of this section.

Content

In media production and publishing, content is information and experiences that may provide value for an end-user/audience. Content may be delivered via any medium such as the internet, television, and audio CDs, as well as live events such as conferences and stage performances.

The word "content" is often used colloquially to refer to media, which is erroneous as it instead means the contents of the medium rather than the medium itself. Likewise, the single word "media" and some compound words that include "media" (e.g. multimedia, hypermedia) are instead referring to a type of content. An example of a type of content commonly referred to as a type of media is a "motion picture" referred to as "a film”. The distinction between medium and content is less clear when referring to interactive elements that contain information and are then contained in interactive media, such as dice contained in board games or GUI widgets contained in software.

Content Value

The author, producer or publisher of an original source of information or experiences may or may not be directly responsible for the entire value that they attain as content in a specific context. For example, part of an original article (such as a headline from a news story) may be rendered on another web page displaying the results of a user's search engine query grouped with headlines from other news publications and related advertisements. The value that the original headline has in this group of query results may be very different from the value that it had in its original article.

It is possible for a person to derive their own value from content in ways that the author didn't plan or imagine. User innovation makes it possible for users to develop their own content from existing content.

Not all content requires creative authoring or editing. Through recent technological developments such as mobile phones that can record events anywhere for publishing and converting to potentially reach a global audience on channels such as YouTube, most recorded or transmitted information and experiences can be deemed content.

Content Management

Content management is a set of processes and technologies that support the evolutionary life cycle of digital information. This digital information is often referred to as content or, to be precise, digital content. Digital content may take the form of text, such as documents, multimedia files, such as audio or video files, or any other file type which follows a content lifecycle which requires management.

The Process Of Content Management

The digital content life cycle consists of 6 primary phases:

- Create
- Update
- Publish
- Translate
- Archive
- Retrieve

For example, an instance of digital content is created by one or more authors. Over time that content may be edited. One or more individuals may provide some editorial oversight thereby approving the content for publication. Publishing may take many forms. Publishing may be the act of pushing content out to others, or simply granting digital access rights to certain content to a particular person or group of persons. Later that content may be superseded by another form of content and thus retired or removed from use.

Content management is an inherently collaborative process. It often consists of the following basic roles and responsibilities:

- Creator - responsible for creating and editing content.
- Editor - responsible for tuning the content message and the style of delivery, including translation and localization.
- Publisher - responsible for releasing the content for use.
- Administrator - responsible for managing access permissions to folders and files, usually accomplished by assigning access rights to user groups or roles. Admins may also assist and support users in various ways.
- Consumer, viewer or guest- the person who reads or otherwise takes in content after it is published or shared.

A critical aspect of content management is the ability to manage versions of content as it evolves. Authors and editors often need to restore older versions of edited products due to a process failure or an undesirable series of edits.

Another equally important aspect of content management involves the creation, maintenance, and application of review standards. Each member of the content creation and review process has a unique role and set of responsibilities in the development and/or publication of the content. Each review team member requires clear and concise review standards which must be maintained on an ongoing basis to ensure the long-term consistency and health of the knowledge base.

A content management system is a set of automated processes that may support the following features:

- Import and creation of documents and multimedia material
- Identification of all key users and their roles
- The ability to assign roles and responsibilities to different instances of content categories or types.
- Definition of workflow tasks often coupled with messaging so that content managers are alerted to changes in content.
- The ability to track and manage multiple versions of a single instance of content.
- The ability to publish the content to a repository to support access to the content. Increasingly, the repository is an inherent part of the system, and incorporates enterprise search and retrieval.

Content management systems take the following forms:

- a web content management system is software for web site management - which is often what is implicitly meant by this term
- the work of a newspaper editorial staff organization
- a workflow for article publication
- a document management system
- a single source content management system - where content is stored in chunks within a relational database

Core Requirements

A content management system usually requires the following features:

- Identification of all key users and their content management roles.
- The ability to assign roles and responsibilities to different content categories or types.
- Definition of the content work flow tasks, often coupled with event messaging so that content managers are alerted to changes in content.
- The ability to track and manage multiple versions of a single instance of content.
- The ability to publish the content to a repository to support access to the content. Increasingly, the repository is an inherent part of the system, and incorporates enterprise search and retrieval.
- Some content management systems allow the semantic layer of content to be separated to some extent from its layout. For example the CMS may automatically set the colour, fonts, or emphasis of text.

Digital Asset Management

Digital Asset Management consists of tasks and decisions surrounding ingesting, annotating, cataloguing, storage and retrieval of digital assets, such as digital photographs, animations, videos and music. Digital asset management systems are computer software and/or hardware systems that aid in the process of digital asset management.

The term "Digital Asset Management" (DAM) also refers to the protocol for downloading, renaming, backing up, rating, grouping, archiving, optimizing, maintaining, thinning, and exporting files.

The term "Media Asset Management" (MAM) is sometimes used as a sub-category of "Digital Asset Management", mainly for audio or video content. The more recent concept of Enterprise Content Management (ECM) often describes solutions which address similar features but in a wider range of industries or applications.

Uses Of Digital Asset Management

Many businesses and organizations are adopting Digital Asset Management as a business strategy because managing image, video and other media assets presents unique challenges and requires solutions designed specifically to streamline the acquisition, storage and retrieval of digital media. Effective implementation of a DAM system should reduce the time and cost of content production, maximize the return on investment (ROI) from media assets, bring new products and services to market faster and streamline compliance. This system should be designed in such a way that enables cost-effective optimization of media asset management across an organization.

Digital Asset Management Technical Context

Generally the "asset" being managed is collected and stored in a digital format. There is usually a target version of that referred to as "essence" and is generally the highest resolution and fidelity representation. The asset is detailed by its "metadata". Metadata is the description of the asset and the description depth can vary depending on the needs of the system, designer, or user. Metadata can describe, but is not limited to, the description of: asset content (what is in the package?); the means of encoding/decoding (e.g. JPEG, tar, MPEG 2); provenance (history to point of capture); ownership; rights of access; as well as many others.

There exist some predefined standards and template for metadata such as Dublin Core and PBCore. In cases of systems that contain large size asset essences, such as MPEG 2 and JPEG2000 for the cases of images and video, there are usually related "proxy" copies of the essence. A proxy copy is a lower resolution representation of the essence that can be used as a reference in order to reduce the overall bandwidth requirements of the DAM system infrastructure. It can be generated and retained at the time of ingestion of the asset simultaneous or subsequent to the essence or it can be generated on the fly using transcoders.

Smaller DAM systems are easier to categorize as to content and usage since they would normally operate in a particular operational context. This would hold true for systems attached to audio or video production systems. The key differentiators here are the type of decoders and I/O (input/output) used for the asset ingest, use and outgest. Since the essence (and proxy copies) are described by metadata, the metadata can be used as a guide to the playout decoders, transcoders, and channels as well as a input to access control rules.

This means that essentially the essence can be treated as a non-described storage object except when being accessed for viewing or editing. There is relevance to this when considering the overall design and use of larger implementations. The closer the asset is to the ingest/edit/playout tool, the greater the technical architecture needs to accommodate delivery requirements such as bandwidth, latency, capacity, access control, availability of resources, etc. The further the asset moves into a general storage architecture (e.g. Hierarchical Storage Management [HSM]) the more it can be treated as a general blob (binary large object) that is typically held in the file system, not the database. The impact of this set of needs means that it is possible and reasonable to design larger systems using smaller, more expensive performance systems at the edge of the network where the essence is being used in its intended form and less expensive systems further back for storage and archival. This type of design is an example of Infrastructure Convergence Architecture where the line of business operations technology and IT technologies are dependent on one another for functional and performance (non-functional requirements) requirements.

Types of Digital Asset Management Systems

The following broad categories of digital asset management systems may be distinguished:

- Brand asset management systems, with a focus on facilitation of content re-use within large organizations. Here the content is largely marketing- or sales-related, for example, product imagery, logos, marketing collateral or fonts, to give a few examples.
- Library asset management systems, with a focus on storage and retrieval of large amounts of infrequently changing media assets, for example in video or photo archiving.
- Production asset management systems, with a focus on storage, organization and revision control of frequently changing digital assets, for example in digital media production.
- Digital supply chain services, pushing digital content out to digital retailers (e.g. music, videos and games).

Content Delivery

Content delivery describes the delivery of digital media "content" such as digital audio or digital video or computer software and games over a delivery medium such as broadcasting or the Internet.

Content delivery has two parts:

- delivery of finished content for digital distribution, with its accompanying metadata
- delivery of the end product to the consumer

Specialist networks known as "Content Delivery Networks" have emerged to distribute digital content on the Internet. Alternative technologies for content delivery include peer-to-peer file sharing technologies.

Content can only be delivered if it exists. If it does not exist, several techniques and methods can be used for content creation or content regeneration.

Alternatively, Content Delivery Platforms, which are like hosted Content management systems have emerged to create and syndicate content remotely.

Content Delivery Platform (CDP)

A content delivery platform is a Software as a Service (SaaS) content service, similar to a content management system (CMS), which utilizes embedded software code to deliver Web content.

Instead of the installation of software on client servers, a CDP feeds content through embedded code snippets, typically via JavaScript widget, Flash widget or server-side Ajax (programming).

A CDP is not a CDN (Content delivery network). CDNs are utilized for large Web media and do not depend on embedded software code. A CDP is utilized for all types of web content, even text-based content.

Content Delivery Network (CDN)

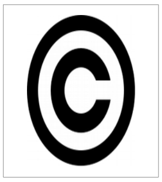
A content delivery network or content distribution network (CDN) is a system of computers networked together across the Internet that cooperate transparently to deliver content (especially large media content) to end users. The first web content based CDNs were Speedera, Sandpiper, Mirror Image and Skycache, followed by Akamai and Digital Island.

CDN nodes are deployed in multiple locations, often over multiple backbones. These nodes cooperate with each other to satisfy requests for content by end users, transparently moving content behind the scenes to optimize the delivery process. Optimization can take the form of reducing bandwidth costs, improving end-user performance, or both.

The number of nodes and servers making up a CDN varies, depending on the architecture, some reaching thousands of nodes with tens of thousands of servers.

Requests for content are intelligently directed to nodes that are optimal in some way. When optimizing for performance, locations that can serve content quickly to the user may be chosen. This may be measured by choosing locations that are the fewest hops or fewest number of network seconds away from the requestor, so as to optimize delivery across local networks. When optimizing for cost, locations that are less expensive to serve from may be chosen instead. Often these two goals tend to align, as servers that are close to the end user sometimes have an advantage in serving costs, perhaps because they are located within the same network as the end user.

Copyright



Copyright is a legal concept, enacted by most governments, giving the creator of an original work exclusive rights to it, usually for a limited time. Generally, it is "the right to copy", but also gives the copyright holder the right to be credited for the work, to determine who may adapt the work to other forms, who may perform the work, who may financially benefit from it, and other, related rights. It is an intellectual property form (like the patent, the trademark, and the trade secret) applicable to any expressible form of an idea or information that is substantive and discrete.

Copyright initially was conceived as a way for government to restrict printing; the contemporary intent of copyright is the promoting the creation of new works by giving authors control of and profit from them. Copy rights have been internationally standardised, lasting between fifty to a hundred years from the creator's death, or a finite period for anonymous or corporate creations; some jurisdictions have required formalities to establishing copyright, most recognize copyright in any completed work, without formal registration. Generally, copyright is enforced as a by civil matter, though some jurisdictions do apply criminal sanctions.

Most jurisdictions recognize copyright limitations, allowing "fair" exceptions to the creator's exclusivity of copyright, and giving users certain rights. The development of digital media and computer network technologies have prompted reinterpretation of these exceptions, introduced new difficulties in enforcing copyright, and inspired additional challenges to copyright law's philosophic basis. Simultaneously, businesses with great economic dependence upon copyright have advocated the extension and expansion of their copy rights, and sought additional legal and technological enforcement.

Copyleft

Copyleft is a play on the word copyright and describes the practice of using copyright law to remove restrictions on distributing copies and modified versions of a work for others and requiring that the same freedoms be preserved in modified versions.

Copyleft is a form of licensing and may be used to modify copyrights for works such as computer software, documents, music, and art. In general, copyright law allows an author to prohibit others from reproducing, adapting, or distributing copies of the author's work. In contrast, an author may, through a copyleft licensing scheme, give every person who receives a copy of a work permission to reproduce, adapt or distribute the work as long as any resulting copies or adaptations are also bound by the same copyleft licensing scheme. A widely used and originating copyleft license is the GNU General Public License. Similar licenses are available through Creative Commons — called Share-alike.

Copyleft may also be characterized as a copyright licensing scheme in which an author surrenders some but not all rights under copyright law. Instead of allowing a work to fall completely into the public domain (where no copyright restrictions are imposed), copyleft allows an author to impose some but not all copyright restrictions on those who want to engage in activities that would otherwise be considered copyright infringement. Under copyleft, copyright infringement may be avoided if the would-be infringer perpetuates the same copyleft scheme. For this reason copyleft licenses are also known as reciprocal licenses.

While the copyright only protects the exclusive rights of the originator by excluding all others under penalty, the Copyleft protects the freedom of all others, practically to use copyrighted works as if there were no copyright law at all, with one exception: Furthermore, copyleft also protects the open access to necessary information (e.g. source code) by enforcing its publication.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.*io*

Design Considerations – Mobile

The New Customer – The Arrival Of The Prosumer

In many ways the design case for the moble browser is less a technical matter than a socio-cultural one. Basically this means that new modes of buying and selling that have emerged in recent years are superceding the classic MBA view of supply and demand.

These new styles of customer engagement are at the core of the success drivers that will ensure our solution's success. In this section we review some of these trends with an eye on incorporating them into the solution design.

New Does Not Mean Loss Leader

It must be added, before we begin to review the trends that have influenced this study and which will continue to influence the development of this design, is that none of them are loss leaders. They have sane economic benefits albeit in some case these may be both new to the reader and, in most cases, counter-intuitive.

User Generated Content (UGC)

User generated content refers to various kinds of media content, publicly available, that are produced by end-users. This includes digital video, blogging, podcasting, news, gossip, research, mobile phone photography and wikis. In addition to these technologies, user generated content may also employ a combination of open source, free software, and flexible licensing or related agreements to further diminish the barriers to collaboration, skill-building and discovery.

User generated content has also been characterized as 'Conversational Media', as opposed to 'Packaged Goods Media' (that is, traditional media). The former is a two-way process in contrast to the one-way distribution of the latter. Conversational or two-way media is a key characteristic of so-called Web 2.0 which encourages the publishing of one's own content and commenting on other people's.

The notion of the passive audience therefore has shifted since the birth of New Media, and an ever-growing number of participatory users are taking advantage of the interactive opportunities, especially on the Internet to create independent content. Grassroots experimentation then generated an innovation in sounds, artists, techniques and associations with audiences which then are being used in mainstream media. The active, participatory and creative audience is prevailing today with relatively accessible media, tools and applications, and its culture is in turn impacting mass media corporations and global audiences.

The Organisation for Economic Co-operation and Development (OECD) has defined three central characteristics for UGC:

1. **Publication requirement:** While UGC could be made by a user and never published online or elsewhere, we focus here on the work that is published in some context, be it on a publicly accessible website or on a page on a social networking site only accessible to a select group of people (eg, fellow university students). This is a useful way to exclude email, two-way instant messages and the like.
2. **Creative effort:** This implies that a certain amount of creative effort was put into

creating the work or adapting existing works to construct a new one; i.e. users must add their own value to the work. The creative effort behind UGC often also has a collaborative element to it, as is the case with websites which users can edit collaboratively. For example, merely copying a portion of a television show and posting it to an online video website (an activity frequently seen on the UGC sites) would not be considered UGC. If a user uploads his/her photographs, however, expresses his/her thoughts in a blog, or creates a new music video, this could be considered UGC. Yet the minimum amount of creative effort is hard to define and depends on the context.
3. **Creation outside of professional routines and practices:** User generated content is generally created outside of professional routines and practices. It often does not have an institutional or a commercial market context. In extreme cases, UGC may be produced by non-professionals without the expectation of profit or remuneration. Motivating factors include: connecting with peers, achieving a certain level of fame, notoriety, or prestige, and the desire to express oneself.

Copy & paste or a link could also be seen as user generated self-expression. The action of linking to a work or copying a work could in itself motivate the creator, express the taste of the person linking or copying. Digg.com, Stumbleupon.com, leaptag.com is a good example where such linkage to work happens. The culmination of such linkages could very well identify the tastes of a person in the community and make that person unique through statistical probabilities.

Types of of user generated content include:

- Discussion boards
- Blogs
- Social networking sites
- News Sites
- Trip planners
- Customer review sites
- Experience or photo sharing sites
- Any other website that offers the opportunity for the consumer to share their knowledge and familiarity with a product or experience
- User generated games

Social Media

Social media is an umbrella term that defines the various activities that integrate technology, social interaction, and the construction of words, pictures, videos and audio. This interaction, and the manner in which information is presented, depends on the varied perspectives and "building" of shared meaning, as people share their stories, and understandings.

Social media can take many different forms, including Internet forums, message boards, weblogs, wikis, podcasts, pictures and video. Technologies include: blogs, picture-sharing, vlogs, wall-postings, email, instant messaging, music-sharing, crowdsourcing, and voice over IP, to name a few. Examples of social media applications are Google Groups (reference, social networking), Wikipedia (reference), MySpace (social networking), Facebook (social networking), Last.fm (personal music), YouTube (social networking and video sharing), Second Life (virtual reality), Flickr (photo sharing), Twitter (social networking and microblogging) and other microblogs are Jaiku and Pownce. Many of these social media services can be integrated via Social network aggregation platforms like Mybloglog, a Yahoo property, and Plaxo.

Social media or social networking (one example of social media) has a number of characteristics that make it fundamentally different from traditional media such as newspapers, television, books, and radio. Primarily, social media depends on interactions between people as the discussion and integration of words builds shared-meaning, using technology as a conduit.

Social media utilities create opportunities for the use of both inductive and deductive logos by its users. Claims or warrants are quickly transitioned into generalizations due to the manner in which shared statements are posted and viewed by all. The speed of communication, breadth, and depth, and ability to see how the words build a case solicits the use of rhetoric. Induction is frequently used as a means to validate or authenticate different user's statement, and words. Rhetoric is an important part of today's language in social media.

Social media is not finite: there is not a set number of pages or hours. The audience can participate in social media by adding comments or even editing the stories themselves. Content in social media can take the form of text, graphics, audio, or video. Several formats can be mixed. Social media is typically available via feeds, enabling users to subscribe via feed readers, and allowing other publishers to create mashups.

Social media signifies a broad spectrum of topics and has several different connotations. In the context of Internet marketing, Social Media refers to a collective group of web properties that are driven by users. For example, blogs, discussion boards, vlogs, video sharing sites and dating sites. Social Media Optimization (SMO) is the process of trying to get ones content more widely distributed across multiple Social Media networks.

Social Media has two important aspects. The first, SMO, refers to on-page tactics through which a webmaster can improve a website for the age of social media. Such optimization includes adding links to services such as Digg, Reddit and Del.icio.us so that their pages can be easily 'saved and submitted' to and for these services.

Social Media Marketing, on the other hand, is an off-page characteristic of Social Media. This includes writing content that is remarkable, unique, and newsworthy. This content can then be marketed by popularizing it or even by creating a "viral" video on YouTube and other video sites. Social Media is about being social so this off-page work can include getting involved in other similar blogs, forums, and niche communities. Search Engine Marketing or SEM involves utilization of all available Social Networking platforms to brand a product using Search Engine Optimization or SEO techniques of communication, to the end consumer.

Mobile Media (Or A Homage To Paul Levinson)

The mobility and portability of media or as Paul Levinson calls it in his book entitled Cellphone "the media-in-motion business" has been a process in the works ever since the "first time someone thought to write on a tablet that could be lifted and hauled – rather than on a cave wall, a cliff face, a monument that usually was stuck in place, more or less forever". Today, mobile media devices such as mobile phones and PDA's are the primary source of portable media from which we can obtain information and communicate with one another.

While mobile phone and PDA's independent technologies and functions may be new and innovative (in relation to changes and improvements in media capabilities in respect to their function what they can do when and where and what they look like, in regard to their size and shape) the need and desire to access and use media devices regardless of where we are in the world has been around for centuries. Indeed Paul Levinson remarks in regard to telephonic communication that it was "intelligence and inventiveness applied to our need to communicate regardless of where we may be, led logically and eventually to telephones that we carry in our pockets". Levinson in his book goes on to state that the book, transistor radio, Kodak camera are also bearers of portable information. And that it is thanks to the printing press that information became available to a mass audience, the reduction in size and portability of the camera allowed people to capture what they saw no matter where they were and the internet meant that people could talk to anyone and use on demand information.

Mobile phones, digital cameras, ipods, walkmans, laptops, PDA, Game Boys and their corresponding media technologies play an increasingly important role in the everyday lives of millions of people world wide. Media can be downloaded onto the device by podcasting or can be streamed over the web. Digital applications include gaming, video, audio, downloadable ring tones and mobizines. A new direction is that of Mobile TV.

Crowdsourcing

Crowdsourcing is a neologism for the act of taking a task traditionally performed by an employee or contractor, and outsourcing it to an undefined, generally large group of people, in the form of an open call. For example, the public may be invited to develop a new technology, carry out a design task, refine an algorithm or help capture, systematize or analyze large amounts of data.

The term has become popular with business authors and journalists as shorthand for the trend of leveraging the mass collaboration enabled by Web 2.0 technologies to achieve business goals.

In some cases the labor is well-compensated, but i most general n other cases the only rewards may be kudos or intellectual satisfaction. Barter agreements are often associated with crowdsourcing. Crowdsourcing may produce solutions from amateurs or volunteers working in their spare time, or from small businesses which were unknown to the initiating organization.

Perceived benefits of crowdsourcing include:

- Problems can be explored at comparatively little cost.
- Payment is by results.
- The organization can tap a wider range of talent than might be present in its own organisation.

The difference between crowdsourcing and ordinary outsourcing is that a task or problem is outsourced to the public, rather than another body. The difference between crowdsourcing and open source is that open source production is a cooperative activity initiated and voluntarily undertaken by members of the public. In crowdsourcing the activity is initiated by a client, and the work may be undertaken on an individual, as well as a group, basis.

Some of the risks of crowdsourcing include:

- Added costs post-completion of a project to bring a project to an acceptable conclusion.
- Increased likelihood that a crowdsourced project will suffer failure due to lack of monetary motivation, too few participants, lower quality of work, lack of personal interest in the project, global language barriers, or difficulty managing a large-scale crowdsourced project.

- No written contracts, non-disclosure agreements, or employee agreements or agreeable terms with crowdsourced employees.
- Difficulties maintaining a working relationship with crowdsourced workers throughout the duration of a project.
- Susceptibility to faulty results caused by targeted, malicious work efforts.

Open Innovation

The central idea behind open innovation is that in a world of widely distributed knowledge, companies cannot afford to rely entirely on their own research, but should instead buy or license processes or inventions (i.e. patents) from other companies. In addition, internal inventions not being used in a firm's business should be taken outside the company (e.g., through licensing, joint ventures, spin-offs). In contrast, closed innovation refers to processes that limit the use of internal knowledge within a company and make little or no use of external knowledge. Some companies promoting open innovation include Procter & Gamble and IBM.

While open source and open innovation might conflict on patents issues, they are not mutually exclusive, as participating companies can donate their patents to an independent organization, put them in a common pool or grant unlimited license use to anybody. Hence some open source initiatives can merge the two concepts, this is the case for instance for IBM with its Eclipse platform which IBM is advocating as a case of open innovation, where competing companies are invited to co-operate inside an open innovation network.

Open innovation needs a different mindset and company culture than traditional or closed innovation.

Closed innovation Principles	Open innovation Principles
The smart people in our field work for us.	Not all the smart people work for us. We need to work with smart people inside and outside our company.
To profit from research and development (R&D), we must discover it, develop it and ship it ourselves.	External R&D can create significant value; internal R&D is needed to claim some portion of that value.
If we discover it ourselves, we will get it to market first.	We don't have to originate the research to profit from it.
The company that gets an innovation to market first will win.	Building a better business model is better than getting to market first.
If we create the most and the best ideas in the industry, we will win.	If we make the best use of internal and external ideas, we will win.
We should control our innovation process, so that our competitors don't profit from our ideas.	We should profit from others' use of our innovation process, and we should buy others' intellectual property (IP) whenever it advances our own business model.

Collaboration, The Commons & Massive Participation

Collaborative methods are processes, behaviors and conversations that relate to collaboration between individuals.[1] These methods specifically aim to increase the success of teams as they engage in collaborative problem solving. Forms, rubrics, charts and graphs are useful in these situations to objectively document personal traits with the goal improving performance in current and future projects.

There are seven rules for all collaboration:

- Look for common ground: find shared values, consider shared personal experiences, pay attention to and give feedback, be yourself and expect the same of others, be willing to accept differences in perception and opinions

- Learn about others: consider their perspectives and needs, appeal to the highest motives, let others express themselves freely

- Critique results, not people: do not waste time on personal hostility, make other people feel good, avoid criticism and put downs

- Give and get respect: show respect for others opinions, be considerate and friendly, put yourself in the other person's shoes, be responsive to emotions, speak with confidence but remain tactful

- Proceed slowly: present one idea at a time, check for understanding and acceptance of each idea before moving on to the next. Speak in an organized and logical sequence.

- Be explicit and clear: share your ideas and feelings, pay attention to nonverbal communication, speak clearly and make eye contact, select words that have meaning for your listeners

- Remember the five "Cs" of communication: clarity, completeness, conciseness, concreteness, and correctness

  (Source: Spence, Muneera U. "Graphic Design: Collaborative Processes = Understanding Self and Others." (lecture) Art 325: Collaborative Processes. Fairbanks Hall, Oregon State University, Corvallis, Oregon. 13 Apr. 2006.)

Examples of products created by means of commons-based peer production include Linux, a computer operating system; Slashdot, a news and announcements website; Kuro5hin, a discussion site for technology and culture; Wikipedia, an online encyclopedia; and Clickworkers, a citizen science program.

Several unexpected but foreseeable outgrowths have been:

- Customization/Specialization. With free and open source software small groups are capable to customize a large project to specific needs.
- Immortality. Once code is released under a copyleft free software license the genie cannot be put back into the bottle.
- Cross-fertilization. Experts in a field can work on more than one project with no legal hassles.
- Technology Revisions: A core technology gives rise to new implementations of existing projects.
- Technology Clustering: Groups of products tend to cluster around a core set of technology and integrate with one another.

Massively Distributed Collaboration For Content Creation (MDC3)

A notoriously difficult concept to define albeit absolutely critical in understadning the potential of the mobile browser as content factory.

An attempt at defining massively distributed collaboration for content creation:

- **Collaboration** - The key to this is meaningful and significant feedback: without it, signals may be sent and received, but there is no collaboration. So a feedback-less television show or ad campaign or education session would not be covered, by the definition.
- **Distributed** - The idea here is the Internet multi-nodal model, the one which DARPA legendarily liked because the foreign bombs couldn't take it out easily: an information system which is greatly centralized, like the French Minitel or most traditional corporate and government information systems, would not be covered by the definition.
- **Massively** - The scale of the Internet, and of globalization and other enormous information dissemination and sharing projects now under way, gets folded into the definition here. Systems deliberately restricted to small-scale use, relative to other systems which are not, would not be covered by the definition. "Distributed collaboration for content creation" in your household or within your (small) academic department or corporation probably would not qualify: the idea here is the "mass" scale of the system -> bigger is better, here.
- **Content Creation** - "Massively distributed collaboration" can be undertaken for many purposes: to build a dam in China, to sell automobiles nationwide, to distribute medical assistance in sub-Saharan Africa. But if the effort is one-shot, the automobiles get sold and there is nothing left over thereafter, it is not "content creation". The content, the data-plus-intelligence which yields "information", is the point. Certain Global efforts might be admirable, in other words, and require collaboration which is massive and distributed, but the key idea behind the definition as it is used here is that such efforts should concern information systems: meaningful data which can be archived, and searched & retrieved, and used over and over again.

The term “massively distributed collaboration” was coined by Mitchell Kapor, in a presentation at UC Berkeley on 2005-11-09, to describe an emerging activity of wikis and electronic mailing lists and blogs and other content-creating virtual communities online. Kapor said, in the introduction to his talk:

> *"The sudden and unexpected importance of the Wikipedia, a free online encyclopedia created by tens of thousands of volunteers and coordinated in a deeply decentralized fashion, represents a radical new modality of content creation by massively distributed collaboration. This talk will discuss the unique principles and values which have enabled the Wikipedia community to succeed and will examine the intriguing prospects for application of these methods to a broad spectrum of intellectual endeavors"*

MDC3 Tools

- Wikis — massively distributed collaboration is not the only purpose served by a wiki, such as Wikipedia, but it can be the central purpose and at least is a by-product, insofar as participants are assembling a body of data or information which can be re-used by themselves and used by others.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited*****.io***

- eConferences / Listservs — the archives, of econferences and listservs some of which have been in operation since the early 1990s, are the best records we possess of the establishment and development of digital information techniques, the Gutenberg Bible of the Digital Era — one hopes such archives will be preserved.

- Gaming Boards / Product Forums — a vast range of topics gets discussed, in some of these an entire younger generation obtains and develops much of its peer-group collective knowledge on these things.

- Blogs — as for eConferences so for blogs, archives of these are at least as important as the realtime interactions they host: the best records of the Age of Incunabula of the Digital Era may well become blog archives

- Wi-Fi — any tool which extends the "distribution", "massively", to enable the "collaboration" — the increasing transparency (XeroxPARC term) of the technology is its basic enabling device, letting the machinery blend back into the woodwork just as the telephone did.

- Mobile Phones — another tool extending "distribution", "massively" — this one directly offering interactivity to enable "collaboration" such as improvements in screen technology which more enable passive Infotainment etc. applications.

- Virtual Communities — another digital tool directly aiming at "collaboration" in "content creation" — as passive Infotainment, one-to-one email or local & restricted & possibly centralized systems. A good argument can be made that the walled garden of our solution should be viewed and described as a highly centralised virtual community allowing viral interactions inside the walls for members.

MDC3 Applications

A small list would include:

- Education (inc distance education, continuing education / career retraining and lifelong learning)
- Research
- Corporations (see Business Week Magazine, US edition November 21 2005 — "Best Practices, Smart Ways to Use the Web, Companies That Get It" — they provide a list and almost all of their praise goes to "massively distributed collaboration" applications)
- Communities
- Emergency relief
- Literature
- Music
- Political Action
- Legislation

User Innovation Toolkits

A design criteria to enable MDC3 (suggested by Eric Von Hippel) would have certain requirements, being:

- Learning by trial-and-error
- An appropriate solution space

- A user friendly toolkit
- Commonly used modules
- Result easily created by producer

Learning By Trial-And-Error

It is important that the user is able to go through complete trial-and-error cycles when designing the product. This allows the users to see the consequences of the design choices they make, and thereby decide more precisely what they really want. Trial-and-error has been shown by research to be the way that most problem solving is done.

An Appropriate Solution Space

A solution space is defined by the flexibility in which the producer can produce the desired result. Any production process has a set of limiting factors, and these factors define the solution space. If the solution space is small, the chance of user innovations are small.

A User Friendly Toolkit

The process must be available to the users so that they can use the skills and languages they already know. This frees the users from learning the different design-specific skills and languages associated with manufacturing.

Commonly Used Modules

Custom designs are seldom made up of unique parts, but instead share a set of standard modules. Therefore a library of standard modules should be available to the user. This allows the user to focus on the unique parts that are truly important.

Results Easily Created

The result from the process must be easily converted into the language needed for the production system, and be without error. If the result of the process must be manually translated much of the effect of the toolkit may be lost.

The Prosumer

Prosumer is a term formed by contracting either the word producer or professional with the word consumer. The term has taken on multiple conflicting meanings: the business sector sees the prosumer (professional–consumer) as a market segment, whereas economists see the prosumer (producer–consumer) as having greater independence from the mainstream economy.

Producer & Consumer

Marshall McLuhan and Barrington Nevitt suggested in their book Take Today that with electric technology, the consumer would become a producer. In the 1980 book, The Third Wave, futurologist Alvin Toffler coined the term "prosumer" when he predicted that the role of producers and consumers would begin to blur and merge (even though he described it in his book Future Shock from 1970). Toffler envisioned a highly saturated marketplace as mass production of standardized products began to satisfy basic consumer demands. To continue growing profit, businesses would initiate a process of mass customization, that is the mass production of highly customized products.

However, to reach a high degree of customization, consumers would have to take part in the production process especially in specifying design requirements. In a sense, this is merely an extension or broadening of the kind of relationship that many affluent clients have had with professionals like architects for many decades.

Toffler has extended these and many other ideas well into the 21st-century. Along with recently published works such as Revolutionary Wealth (2006), we can recognize and assess both the concept and fact of the prosumer as it is seen and felt on a worldwide scale. That these concepts are having global impact and reach, however, can be measured in part by noting in particular, Toffler's popularity in China. Discussing some of these issues with Newt Gingrich on C-Span's After Words program in June 2006, Toffler mentioned that The Third Wave is the second ranked bestseller of all time in China, just behind a work by Mao Zedong.

Don Tapscott more fully elaborated on the concept in his 1995 book The Digital Economy calling it "Prosumption."

More recently, The Cluetrain Manifesto noted that "markets are conversations" with the new economy moving from passive consumers ... to active prosumers.2 For instance, Amazon.com emerged as an ecommerce leader -- partially due to its ability to construct customer relations as conversations rather than simple, one-time sales. Amazon supports exchange of information among customers; it provides spaces for customers to add to the site, in the form of reviews.

However, mass customization has not taken place in most areas of the economy. Most consumption continues to be passive as critics of television, recorded music, and fast food would argue. Indeed, people are generally uninterested in going to the effort of customizing the myriad products that comprise modern consumer culture. In The Paradox of Choice: Why More is Less, Barry Schwartz argues that diminishing returns from a confusing abundance of consumer choice is producing stress and dissatisfaction. Still, one key area of high-customization is taking place: highly involved hobbyists.

The Professional Consumer

With customization focused on leisure pursuits, Toffler's initial combination has been largely supplanted by a second pair of blurring roles: that of the professional and consumer. In particular, hobbyists have become ever-more demanding in the pursuits of their hobbies, often rising above the level of dilettante (an amateur, someone who dabbles in a field out of casual interest rather than as a profession or serious interest) to the point of commanding skills equal to that of professionals. Key examples of such hobbies are:

- home improvement as illustrated by the rise of hardware stores such as Home Depot
- cooking as illustrated by creation of entirely new cable television channels entirely dedicated to the culinary arts
- photography as illustrated by still cameras and camcorders that, often, are on a par with equipment used by professionals.

This professional slant of the prosumer term is most common in photography which is a field that highlights prosumer trends. Access to professional-level equipment and skills is made possible by combination of factors such as:

- high disposable incomes by some sectors of the population
- increased leisure time, again, for some sectors of the population

- continuously falling prices of ever more advanced products (especially electronics)
- media geared towards amateurs and hobbyists:
  a. beginning in the 1980s with the advent of desktop publishing, a growing profusion of magazines to satisfy specialized interests
  b. beginning in the mid-1990s with the advent of the Internet, an even wider range of websites with online forums to pool experience
- Pertaining to electronics; are considered to be "on the fence" as a product of lower quality than a professional product, and higher quality (sometimes in the form of bells and whistles) than a consumer product. Some examples include:
  a. Digital camcorders
  b. Still cameras
  c. HDTVs

Non-Corporate Producer & Consumer

A third meaning or usage of prosumer is springing up, especially among some activist groups. That is, the producer and consumer roles are being combined so as to exclude (or at least diminish) the role of the corporate producer; thus, rather than generating higher corporate profits from value-added products, producers would, at best, be reduced to supplying lower-profit commodity inputs. Indeed, the more consumer-oriented prosumer spin is irrelevant to many people with diminished disposable income caused by various economic trends such as globalization, automation, and wealth concentration. Identifiable trends and movements outside of the mainstream economy that have adopted prosumer terminology and techniques include:

- a Do It Yourself (DIY) approach as a means of economic self-sufficiency or simply as a way to survive on diminished income
- the voluntary simplicity movement that seeks personal, social, and environmental goals through prosumer activities such as:
  a. growing one's own food
  b. repairing clothing and appliances rather than buying new items
  c. playing musical instruments rather than listening to recorded music
- use of new media-creation and distribution technologies to foster independent media (see Indymedia); many involved in independent media reject mass culture generated by concentrated corporate media
- self-sufficient barter networks, notably in developing nations, such as Argentina's RGT have adopted the term prosumer

These blurrings of the roles of consumer and producer have their predecessor in the cooperative self-help movements that sprang up during various economic crises e.g. the Great Depression in the 1930s.

The Influence Of The Prosumer On R&D Budgets

A fourth view of the Prosumer is as one who can influence the R&D spend of a company towards a solution that directly benefits them. For example, say you’re a manufacturer of widgets. One of your strategic customers changes their requirements and asks that all their widgets sing. This is an important enough customer that losing them would seriously hurt your bottom line.

Based on their request you direct a portion of your R & D budget to solve their specific problem. By using internal resources, open source initiatives and outside help you’re able to meet the requirements of your customer. While the customer didn’t directly make the changes they did influence your company with their design requirement changes. This arrangement has positive effects for both parties:

For the customer:

- Immediate access to the new technology.
- Requirements specific to their needs.
- Leverage their relationship with the supplier.

For the company:

- Strengthen the relationship with the customer.
- Demonstrates a willingness to keep their customers satisfied.
- The company now has a new feature/product/service they can market to other customers.

As customers continue to demand more of their supplier relationships, this type of Prosumer influence will only increase. Suppliers up to the task will strengthen those existing relationships, build customer loyalty and become more profitable.

Prosumption

In their book Wikinomics: How Mass Collaboration Changes Everything, Don Tapscott and Anthony D. Williams coined the related term of prosumption (production/consumption) to refer to the creation of products and services by the same people who will ultimately use them. Companies and individuals are increasingly utilizing and involving the end-users to develop final products and services. In some instances, end-users are creating products on their own, without the interference or assistance of third-parties (i.e. companies, organizations, etc). For example, Lego Mindstorms allows users to download software from Lego's website so that the users can edit and update software as they wish.

Ubiquitous Computing

One can rightly classify the mobile browser in the context of our solution as a form of ubiquitous computing. An important reference in this regard is “Spontaneous Emergence Model for Pervasive Environments” published by Gaber in November 2007.

Ubiquitous computing is a post-desktop model of human-computer interaction in which information processing has been thoroughly integrated into everyday objects, dedicated appliances and activities. As opposed to the desktop paradigm, in which a single user consciously engages a single device for a specialized purpose, someone "using" ubiquitous computing engages many computational services and systems simultaneously, in the course of ordinary activities and interaction, and may not necessarily even be aware that they are doing so.

Some models of ubiquitous computing include

- Pervasive Computing
- Ambient Intelligence
- Everyware
- Physical Computing
- "The Internet Of Things"
- Haptic Computing

- "Things That Think"

At their core, all models of ubiquitous computing share a vision of small, inexpensive, robust networked processing devices, distributed at all scales throughout everyday life and generally turned to distinctly quotidian ends. For example, a domestic ubiquitous computing environment might interconnect lighting and environmental controls with personal biometric monitors woven into clothing so that illumination and heating conditions in a room might be modulated, continuously and imperceptibly. Another common scenario posits refrigerators "aware" of their suitably-tagged contents, able to both plan a variety of menus from the food actually on hand, and warn users of stale or spoiled food.

Contemporary human-computer interaction models, whether command-line, menu-driven, or GUI-based, are inappropriate and inadequate to the ubiquitous case. This suggests that the "natural" interaction paradigm appropriate to a fully robust ubiquitous computing has yet to emerge - although there is also recognition in the field that in many ways we are already living in an ubicomp world. Contemporary devices that lend some support to this latter idea include mobile phones, digital audio players, radio-frequency identification tags and interactive whiteboards.

Location Location Location!!!

A core tenet of the TAP business case is for the creation of local content by local users for local users. The value of this cannot be overstated enough.

> *"Locative media is many things: A new site for old discussions about the relationship of consciousness to place and other people. A framework within which to actively engage with, critique, and shape a rapid set of technological developments. A context within which to explore new and old models of communication, community and exchange. A name for the ambiguous shape of a rapidly deploying surveillance and control infrastructure." (Source: Transcultural Media Online Reader Introduction by Ben Russel).*

Locative Media is the media of communication bound to a location. This is digital media applied to real places and thus triggering real social interactions. The incentice for participation by users is direct, compelling and urgent.

We may distinguish between location based games, media, services and social networking.

A location-based game is one in which the game play somehow evolves and progresses via a player's location. Thus, location-based games almost always support some kind of localization technology, for example by using satellite positioning like GPS. "Urban gaming" or "Street Games" are typically multi-player location-based games played out on city streets and built up urban environments.

Location-based media delivers multimedia directly to the user of a mobile device dependent upon their location. Location based media allows for the enhancement of any given environment offering explanation, analysis and detailed commentary on what the user is looking at through a combination of video, audio, images and text. The location-aware device can deliver interpretation of cities, parklands, heritage sites, sporting events or any other environment where location based media is required.

Location-based services are information services accessible with mobile devices through the mobile network and utilizing the ability to make use of the location of the mobile device

Some examples of location-based services are:

- Requesting the nearest business or service, such as an ATM or restaurant
- Turn by turn navigation to Evite invitations, or any other address
- Locating people on a map displayed on the mobile phone
- Receiving alerts, such as notification of a sale on gas or warning of a traffic jam

For the us, the vendor, location-based services provide added value by enabling services such as:

- Resource tracking with dynamic distribution. Taxis, service people, rental equipment, doctors, fleet scheduling.
- Resource tracking. Objects without privacy controls, using passive sensors or RF tags, such as packages and train boxcars.
- Finding someone or something. Person by skill (doctor), business directory, navigation, weather, traffic, room schedules, stolen phone, emergency calls.
- Proximity-based notification (push or pull). Targeted advertising, buddy list, common profile matching (dating), automatic airport check-in.
- Proximity-based actuation (push or pull). Payment based upon proximity (EZ pass, toll watch).

Location-based social networking is an extension of physical and/or actual relationships and social interactions in a given area. It has an added dimension of personalising and enabling social activities, such as finding new friends, groups with similiar interests or even matchmaking, dating and sourcing new sexual partners.

Mobile Telephony – SMS

What Is It?

Short Message Service (SMS) is a service available on most digital mobile phones, other mobile devices (e.g. a Pocket PC, or occasionally even desktop computers) and some fixed phones that permit the sending of short messages between mobile phones, other handheld devices and even landline telephones. The terms text messaging, text messages, more colloquially SMSes, texts, or even txts and its variants are more commonly used in North America, the UK, Spain and the Philippines, while most other countries prefer the term SMS.

Text messages are also often used to interact with automated systems, such as ordering products and services for mobile phones, or participating in contests. There are many services available on the Internet that allows users to send text messages free of charge.

History

As with most other services and modules of functionality of the GSM system, no individual can claim the parenthood of SMS. It might be worthwhile to note this, since such attempts may still be seen - also from people that never took part in the GSM work on SMS. The idea of adding text messaging to the services of mobile users was latent in many communities of mobile communication services at the beginning of the 1980s. Experts from several of those communities contributed in the discussions on which should be the GSM services. Most thought of SMS as a means to alert the individual mobile user, e.g. on incoming voice mail, whereas others had more sophisticated applications in their minds, e.g. telemetry. However, few believed that SMS would be used as a means for sending text messages from one mobile user to another.

As early as February 1985, after having already been discussed in GSM subgroup WP3, chaired by J. Audestad, SMS was considered in the main GSM group as a possible service for the new digital cellular system. In GSM document 'Services and Facilities to be provided in the GSM System' (GSM Doc 28/85 rev2, June 1985), both mobile originated and mobile terminated, including point-to-point and point-to-multipoint, short messages appear on the table of GSM teleservices.

The discussions on the GSM services were then concluded in the recommendation GSM 02.03 'TeleServices supported by a GSM PLMN'. Here a rudimentary description of the three services was given: 1) Short message Mobile Terminated / Point-to-Point, 2) Short message Mobile Originated / Point-to-Point and 3) Short message Cell Broadcast. This was handed over to a new GSM body called IDEG (the Implementation of Data and Telematic Services Experts Group), which had its kickoff in May 1987 under the chairmanship of Friedhelm Hillebrand. The technical standard known today was largely created by IDEG (later WP4) as the two recommendations GSM 03.40 (the two point-to-point services merged together) and GSM 03.41 (cell broadcast).

The first commercial SMS message was sent over the Vodafone GSM network in the United Kingdom on 3 December 1992, from Neil Papworth of Sema Group (using a personal computer) to Richard Jarvis of Vodafone (using an Orbitel 901 handset). The text of the message was "Merry Christmas". The first SMS typed on a GSM phone is claimed to have been sent by Riku Pihkonen, an engineer student at Nokia, in 1993.

Initial growth was slow, with customers in 1995 sending on average only 0.4 messages per GSM customer per month. One factor in the slow takeup of SMS was that operators were slow to set up charging systems, especially for prepaid subscribers, and eliminate billing fraud, which was possible by changing SMSC settings on individual handsets to use the SMSCs of other operators. Over time, this issue was eliminated by switch billing instead of billing at the SMSC and by new features within SMSCs to allow blocking of foreign mobile users sending messages through it. An example of a company that innovate in this subject is OsinetS.A.. By the end of 2000, the average number of messages per user reached 35.

It is also alleged that the fact that roaming customers, in the early days, rarely received bills for their SMSs after holidays abroad had a boost on text messaging as an alternative to voice calls.

SMS was originally designed as part of GSM, but is now available on a wide range of networks, including 3G networks. However, not all text messaging systems use SMS, and some notable alternate implementations of the concept include J-Phone's SkyMail and NTT Docomo's Short Mail, both in Japan. E-mail messaging from phones, as popularized by NTT Docomo's i-mode and the RIM BlackBerry, also typically use standard mail protocols such as SMTP over TCP/IP.

Technical Details

The Short Message Service - Point-to-Point (SMS-PP) is defined in GSM recommendation 03.40. GSM 03.41 defines the Short Message Service - Cell Broadcast (SMS-CB) which allows messages (advertising, public information, etc.) to be broadcast to all mobile users in a specified geographical area.

Messages are sent to a Short Message Service Centre (SMSC), which provides a store-and-forward mechanism. It attempts to send messages to their recipients. If a recipient is not reachable, the SMSC queues the message for later retry. Some SMSCs also provide a "forward and forget" option where transmission is tried only once. Both Mobile Terminated (MT), for messages sent to a mobile handset, and Mobile Originating (MO), for those that are sent from the mobile handset, operations are supported. Message delivery is best effort, so there are no guarantees that a message will actually be delivered to its recipient and delay or complete loss of a message is not uncommon, particularly when sending between networks. Users may choose to request delivery reports, which can provide positive confirmation that the message has reached the intended recipient, but notifications for failed deliveries are unreliable at best.

Transmission of the short messages between SMSC and phone can be done through different protocols such as SS7 within the standard GSM MAP framework or TCP/IP within the same standard. Messages are sent with the additional MAP operation forward_short_message, whose payload length is limited by the constraints of the signalling protocol to precisely 140 bytes (140 bytes = 140 * 8 bits = 1120 bits). In practice, this translates to either 160 7-bit characters, 140 8-bit characters, or 70 16-bit characters. Characters in languages such as Arabic, Chinese, Korean, Japanese or Slavic languages (e.g. Russian) must be encoded using the 16-bit UCS-2 character encoding (see Unicode). Routing data and other metadata is additional to the payload size.

Larger content (known as long SMS or concatenated SMS) can be sent segmented over multiple messages, in which case each message will start with a user data header (UDH) containing segmentation information. Since UDH is inside the payload, the number of characters per segment is lower: 153 for 7-bit encoding, 134 for 8-bit encoding and 67 for 16-

bit encoding. The receiving phone is then responsible for reassembling the message and presenting it to the user as one long message. While the standard theoretically permits up to 255 segments, 6 to 8 segment messages are the practical maximum, and long messages are billed as equivalent to multiple SMS messages.

Short messages can also be used to send binary content such as ringtones or logos, as well as OTA programming or configuration data. Such uses are a vendor-specific extension of the GSM specification and there are multiple competing standards, although Nokia's Smart Messaging is by far the most common.

The SMS specification has defined a way for external Terminal Equipment, such as a PC or Pocket PC, to control the SMS functions of a mobile phone. The connection between the Terminal Equipment and the mobile phone can be realized with a serial cable, a Bluetooth link, an infrared link, etc. The interface protocol is based on AT commands. Common AT commands include AT+CMGS (send message), AT+CMSS (send message from storage), AT+CMGL (list messages) and AT+CMGR (read message).

Some service providers offer the ability to send messages to land line telephones regardless of their capability of receiving text messages by automatically phoning the recipient and reading the message aloud using a speech synthesizer along with the number of the sender.

Today, SMS is also used for machine-to-machine communication. For instance, there is an LED display machine controlled by SMS, and some vehicle tracking companies like ESITrack use SMS for their data transport or telemetry needs. SMS usage for these purposes are slowly being superseded by GPRS services due to their lower overall costs.

Premium Content

SMS is widely used for delivering digital content such as news alerts, financial information, logos and ringtones. Such messages are also known as premium-rated short messages (PSMS). The subscribers are charged extra for receiving this premium content, and the amount is typically divided between the mobile network operator and the value added service provider (VASP) either through revenue share or a fixed transport fee.

Premium short messages are increasingly being used for "real-world" services. For example, some vending machines now allow payment by sending a premium-rated short message, so that the cost of the item bought is added to the user's phone bill or subtracted from the user's prepaid credits. Recently, premium-messaging companies have come under fire from consumer groups due to a large number of consumers racking up huge phone bills. Some mobile networks now require users to call their provider to enable premium messages from reaching their handset.

A new type of 'free premium' or 'hybrid premium' content has emerged with the launch of text-service websites. These sites allow registered users to receive free text messages when items they are interested go on sale, or when new items are introduced.

Popularity

Short message services are developing very rapidly throughout the world. In 2000, just 17 billion SMS messages were sent; in 2001, the number was up to 250 billion, and 500 billion SMS messages in 2004. At an average cost of USD 0.10 per message, this generates revenues in excess of $50 billion for mobile telephone operators and represents close to 100 text messages for every person in the world.

SMS is particularly popular in Europe, Asia (excluding Japan; see below), Australia and New Zealand. Popularity has grown to a sufficient extent that the term texting (used as a verb meaning the act of mobile phone users sending short messages back and forth) has entered the common lexicon. In China, SMS is very popular, and has brought service providers significant profit (18 billion short messages were sent in 2001 [2]). It is also a very influential and powerful tool in The Philippines, where the average user sends 10-12 text messages a day. The Philippines alone sends on the average 400 million text messages a day, more than the annual average SMS volume of countries in Europe, and even China. SMS is hugely popular in India, where youngsters often exchange lots of text messages, and companies provide alerts, infotainment, news, cricket scores update, railway/airline booking, mobile billing, and banking services on SMS. In India, metropolitan media outlets often take real-time polls and audience opinion through SMS, via reserved 4-digit numbers that redirect the information to the respective aforementioned outlets based on designated prefix codes.

Short messages are particularly popular amongst young urbanites. In many markets, the service is comparatively cheap. For example, in Australia a message typically costs between AUD 0.20 and AUD 0.25 to send (some pre-paid services charge AUD 0.01 between their own phones), compared with a voice call, which costs somewhere between AUD 0.40 and AUD 2.00 per minute (commonly charged in half-minute blocks). Despite the low cost to the consumer, the service is enormously profitable to the service providers. At a typical length of only 190 bytes (incl. protocol overhead), more than 350 of these messages per minute can be transmitted at the same data rate as a usual voice call (9 kbit/s).

Text messaging has become so popular that advertising agencies and advertisers are now jumping into the text message business. Services that provide bulk text message sending are also becoming a popular way for clubs, associations, and advertisers to quickly reach a group of opt-in subscribers. This advertising has proven to be extremely effective, but some insiders worry that advertisers may abuse the power of mobile marketing and it will someday be considered spam.

Morse Code

A few widely publicised speed contests have been held between expert Morse code operators and expert SMS users. Several mobile phones have Morse code ring tones and alert messages. For example, many Nokia mobile phones have an option to beep "S M S" in Morse code when it receives a short message. Some of these phones could also play the Nokia slogan "Connecting people" in morse code as a message tone. There are third-party applications available for some mobile phones that allow Morse input for short messages

Spam

In 2002, an increasing trend towards spamming mobile phone users through SMS prompted cellular service carriers to take steps against the practice, before it became a widespread problem. No major spamming incidents involving SMS had been reported as of March 2007, but the existence of mobile-phone spam has been noted by industry watchdogs, including Consumer Reports magazine and the Utility Consumers' Action Network (UCAN). In 2005, UCAN brought a case against Sprint for spamming its customers and charging $0.10 per text message. The case was settled in 2006 with Sprint agreeing not to send customers Sprint advertisements via SMS.

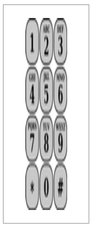
SMS expert LogicaCMG reported a new type of SMS-malice at the end of 2006, noting the first instances of SMiShing (a cousin to email phishing scams). In SMiShing, users receive

SMS messages posing to be from a company, enticing users to phone premium rate numbers, or reply with personal information.

Utility

SMS has also given people instant access to a wealth of information. Services like 82ASK and Any Question Answered in the UK have used the SMS model to enable rapid response to mobile consumers' questions, using on-call teams of experts and researchers.

Text Speak



The small phone keypad caused a number of adaptations of spelling, as in the phrase "txt msg", or use of CamelCase, such as in "ThisIsVeryCool". To avoid the even more limited message lengths allowed when using Cyrillic or Greek letters, speakers of languages written in those alphabets often use the Latin alphabet for their own language.

Historically, this language developed out of shorthand used in Bulletin Board Systems and later in Internet chatrooms, where users would abbreviate some words to allow a response to be typed more quickly. However, this became much more pronounced in SMS, where mobile phone users don't generally have access to a QWERTY keyboard as computer users did, more effort is required to type each character, and there is a limit on the number of characters that may be sent.

In Mandarin Chinese, numbers that sound similar to words are used in place of those words. For example, the numbers 520 in Chinese ("wu er ling") sound like the words for "I love you" ("wo ai ni"). The sequence 748 ("qi si ba") sounds like the curse for "drop dead".

Predictive text software that attempts to guess words (AOL/Tegic's T9 as well as iTAP) or letters (Eatoni's LetterWise) reduces the labour of time-consuming input. This makes abbreviations not only less necessary, but slower to type than regular words which are in the software's dictionary. However it does make the messages longer, often requiring the text message to be sent in multiple parts and therefore costing more to send.

Website portals such as transl8it have supported a community of users to help standardize this text speak by allowing users to submit translations, staking claim with their user handle, or to submit top messages and guess the lingo phrases. The international popularity of this portal resulted in late 2005 the publishing of the transl8it! dxNRE & glosRE (dictionary & glossary) as the worlds first, and most complete, SMS and text lingo book.

Social Impact Of SMS

SMS has caused subtle but interesting changes in society and language since it became popular. Newsworthy events include (in chronological order):

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

Academic Impact

- In December 2002, a cheating scheme was uncovered during final-exam week at the University of Maryland, College Park. A dozen students were caught cheating on an accounting exam through the use of text messages on their mobile phones.
- In December 2002, Hitotsubashi University in Japan failed 26 students for receiving e-mailed exam answers on their mobile phones.
- Using text language is becoming an increasing practice in classes and exams.

Criminal Impact

- In October 2001, a Filipino immigrant living in Belgium was arrested by police after a friend sent him a joke short message pretending to be Osama bin Laden. The message read "I was wondering if I can stay with you for a couple of days. Everybody's so angry with me. I really need a friend. Yours truly, Osama bin Laden."
- In June 2003, a British company developed a program called Fortress SMS for Symbian phones which used 128 bit Rijndael encryption to protect SMS messages.
- In January 2004, cult member Sara Svensson confessed to having murdered the wife of pastor Helge Fossmo and having shot his lover's husband Daniel Linde in Knutby, Sweden. She said that she had acted on anonymous text messages that the pastor had forwarded to her.
- In June 2004, a British punk rock fan was questioned by police regarding a text message sent to a wrong number containing lyrics from "Tommy Gun" by The Clash: "How about this for Tommy Gun? OK - so let's agree about the price and make it one jet airliner for 10 prisoners."
- In July 2004, the police in Tilburg, the Netherlands, started an experiment in which people could register for a short message service. The police would send a message to ask citizens to be vigilant when a burglar was on the loose or a child was missing in their neighbourhood. Several thieves have been caught and children found using the "SMS Alerts". The service has been expanding rapidly to other cities.
- On August 14, 2005, there was a hoax involved in the Helios Airways Flight 522 plane crash. News media widely reported that shortly before the crash a passenger sent a short message indicating that the pilot had become blue in the face, or roughly translated as "The pilot is dead. Farewell, my cousin, here we're frozen." Police later arrested Nektarios-Sotirios Voutas, a 32 year-old private employee from Thessaloniki who admitted that he had made up the story and given several interviews in order to get attention.
- In December 2005, ChinaTechNews.com reported that China's Beijing police detained nine suspects who were members of an illegal wireless short message-sending organization called "Xiao Hai". Local media reported that the suspects included a person surnamed Zou who had been involved in organizing homosexual prostitution, and Wang Wenbin who police say is guilty of bank fraud.
- In February 2005, an Australian company by the name of theSMSzone.com launched a controversial SMS spoofing service allowing messages to be masked, anonymous, and thus totally unidentifiable. This facilitates spam, mobile fraud and defamation, among other things.
- In December 2005 in Australia, text messaging was cited for helping to incite the 2005 Cronulla riots. The SMS messages assisted in mobilising about 5,000 white Australians to engage in violence against those of Middle Eastern origin. In response, some Australians have called for the use of text messaging (or any other electronic means) to incite a riot to be treated as an aggravating circumstance and thus punished more harshly than other forms of incitement.

Political Impact

- In January 2001, Joseph Estrada was forced to resign from the post of president of the Philippines. The popular campaign against him was widely reported to have been co-ordinated with SMS chain letters.
- In the wake of the 2004 Madrid train bombings, SMS was used to garner support for large protest rallies.
- During the 2004 US Democratic and Republican National Conventions, protestors used an SMS based organizing tool called TXTmob.
- During the 2004 Philippine presidential elections, short message was a popular form of electoral campaigning for and against candidates such as incumbent president Gloria Macapagal-Arroyo and main contender Fernando Poe, Jr.
- In March of 2005, SMS was one of the communications forms used to garner support for the Lebanese political rallies.
- French national police spokesman, Patrick Hamon, was quoted in the Wall Street Journal as saying that during the 2005 civil unrest in Franceyouths in individual neighborhoods were communicating by cellphone text messages, online blogs, and/or email — arranging meetings and warning each other about police operations.
- The Islamic Republic of Iran disabled their nationwide SMS network during the 2005 Iranian Presidential elections in which Mahmoud Ahmadinejad was elected President. Some Western commentators have suggested that this was orchestrated to help get Ahmadinejad elected and to quell political uprising.
- Political organisations such as Cymru X, the Plaid Cymru youth wing, and the Young Scots for Independence, the youth wing of the Scottish National Party, have used a "text referendum" to gain public support and raise the profile of their respective causes. The YSI are currently running text referenda on Scottish independence, nuclear weapons, and a St Andrew's Day public holiday.
- In 2006, the Scottish Socialist Party initiated a campaign for people to text the First minister Jack McConnell to demonstrate their support for free school meals.
- SMS messages were used by Chinese nationalists to rapidly spread word of the time and location of demonstrations during the 2005 anti-Japanese demonstrations. At the time, it was one of the few electronic media in China that was not subject to direct government monitoring.

Social Development

- In July 2001, Malaysia's government decreed that an Islamic traditional divorce (which consists of saying "I divorce you" three times in succession) was not valid if sent by short message.
- In 2003, a Malaysian court ruled that, under Sharia law, a man may divorce his wife via text messaging as long as the message was clear and unequivocal.
- In 2003, 2500 employees of the British Amulet Group were fired via a text message to their mobile phone. A similar, widely reported incident occurred in Cardiff, Wales in July 2006.
- In August 2005, an SMS chat sculpture was installed at the annual diploma exhibition of Dresden's University of Art HfBK. The artist Matthias Haase explores today's means of social interaction. Visitors may participate in the artwork by sending a text message to the sculpture, which projects the message onto a screen.
- During Hurricane Katrina in August 2005, many residents were unable to make contact with relatives/friends using traditional landline phones. Via SMS they could communicate with each other when the network worked.[citation needed]

- In November 2006, New Zealand Qualifications Authority approved the move that allowed students of secondary schools to use mobile phone text in the end of the year exam papers.
- In November 2006, Britney Spears reportedly used text messaging to tell her husband Kevin Federline that she is filing for a divorce, however the official divorce filing only occurred the day after the text message was sent. The story was reported by various news media outlets.
- Guinness Book of World records has a world record for text message, currently held by Ang Chuang Yang of Singapore. Ms. Ang keyed in the official text messaging sentence, as established by Guinness ("The razor-toothed piranhas of the genera Serrasalmus and Pygocentrus are the most ferocious freshwater fish in the world. In reality they seldom attack a human."), in 41.52 seconds.

Mobile Telephony – MMS

What Is It?

Multimedia Messaging Service (MMS) is a standard for a telephony messaging systems that allow sending messages that includes multimedia objects (images, audio, video, rich text) and not just text messages as in Short message service (SMS). It is mainly deployed in cellular networks along with other messaging systems like SMS, Mobile Instant Messaging and Mobile E-Mail. Its main standardization effort is done by 3GPP, 3GPP2 and Open Mobile Alliance (OMA).

Basics

MMS is the evolution of Short Message Service (SMS) (SMS is a text-only messaging technology for mobile networks). With MMS, a mobile device is no longer confined to text-only messages. It can send and receive multimedia messages such as graphics, video and audio clips, and so on. It has been designed to work with mobile packet data services such as GPRS and 1x/EVDO.

Methods Of Delivery

There are two modes of delivery in MMS: immediate or deferred:

- Immediate delivery: When the MMS client on the mobile phone receives the MMS notification, it then immediately (without user intervention or knowledge) retrieves the MMS message from the Multimedia Messaging Service Center (MMSC) that sent the notification. After retrieval, the subscriber is alerted to the presence of a newly arrived MMS message.
- Deferred delivery: The MMS client alerts the subscriber that an MMS message is available, and allows the subscriber to choose if and when to retrieve the MMS message.

As with the MMS submission, the MMS retrieval request, whether immediate or deferred, occurs with an HTTP request. The MMSC responds by transmitting the MMS message in an HTTP response to the MMS client, after which the subscriber is finally alerted that the MMS message is available.

The essential difference between immediate and deferred delivery is that the former hides the network latencies from the subscriber, while the latter does not. Immediate or deferred delivery are handset dependent modes, which means that the handset manufacturer can provide the handset in one mode or the other or let the user decide his preference.

Application

MMS-enabled mobile phones enable subscribers to compose and send messages with one or more multimedia parts. Multimedia parts may include text, image, audio and video. These content types should conform to the MMS Standards. For example your phone can send an MPEG-4 video in AVI format, but the other party who is receiving the MMS may not be able to interpret it. To avoid this, all mobiles should follow the standards defined by OMA. Mobile phones with built-in or attached cameras, or with built-in MP3 players are very likely to also have an MMS messaging client -- a software program that interacts with the mobile subscriber to compose, address, send, receive, and view MMS messages.

History

MMS was originally developed within the Third-Generation Partnership Program (3GPP), a standards organization focused on standards for the UMTS/GSM networks.

Since then, MMS has been deployed worldwide and across both GSM/GPRS and CDMA networks.

MMS has also been standardized within the Third-Generation Partnership Program 2 (3GPP2), a standards organization focused on specifications for the CDMA networks.

As with most 3GPP standards, the MMS standards have three stages:

- Stage 1 - Requirements (3GPP TS 22.140)
- Stage 2 - System Functions (3GPP TS 23.140)
- Stage 3 - Technical Realizations

Both 3GPP and 3GPP2 have delegated the development of the Stage 3 Technical Realizations to the OMA, a standards organization focused on specifications for the mobile wireless networks.

Challenges Faced By MMS

There are some interesting challenges with MMS that do not exist with SMS:

- Content adaptation: Multimedia content created by one brand of MMS phone may not be entirely compatible with the capabilities of the recipients' MMS phone. In the MMS architecture, the recipient MMSC is responsible for providing for content adaptation (e.g., image resizing, audio codec transcoding, etc.), if this feature is enabled by the mobile network operator. When content adaptation is supported by a network operator, its MMS subscribers enjoy compatibility with a larger network of MMS users than would otherwise be available.
- Distribution lists: Current MMS specifications do not include distribution lists nor methods by which large numbers of recipients can be conveniently addressed, particularly by content providers, called Value Added Service Providers (VASPs) in 3GPP. Since most SMSC vendors have adopted FTP as an ad-hoc method by which large distribution lists are transferred to the SMSC prior to being used in a bulk-messaging SMS submission, it is expected that MMSC vendors will also likely to adopt FTP.
- Bulk messaging: The flow of peer-to-peer MMS messaging involves several over-the-air transactions that become inefficient when MMS is used to send messages to large numbers of subscribers, as is typically the case for VASPs. For example, when one MMS message is submitted to a very large number of recipients, it is possible to receive a delivery report and read-reply report for each and every recipient. Future MMS specification work is likely to optimize and reduce the transactional overhead for the bulk-messaging case.
- Handset Configuration: Unlike SMS, MMS requires a number of handset parameters to be set. Poor handset configuration is often blamed as the first point of failure for many users. Service settings are sometimes preconfigured on the handset, but mobile operators are now looking at new device management technologies as a means of

delivering the necessary settings for data services (MMS, WAP, etc.) via over-the-air programming (OTA).
- WAP Push: Few mobile network operators offer direct connectivity to their MMSCs for content providers. This has resulted in many content providers using WAP push as the only method available to deliver 'rich content' to mobile handsets. WAP push enables 'rich content' to be delivered to a handset by specifying the URL (via binary SMS) of a pre-compiled MMS hosted on a content provider's web server. A downside of WAP push is that from a billing perspective this content is typically billed at data rates rather than as an MMS. These charges can be significant and result in 'bill shock' for consumers.

MMS should not be confused with Enhanced Messaging Service (EMS), which is simply Short Message Service (SMS) with additional payload capabilities, allowing a mobile phone to send and receive messages that have special text formatting (such as bold or color), animations, pictures, icons, sound effects, and special ring tones.

Protocols

- EAIF — Nokia's External Application Interface
- MM1 — the 3GPP interface between MMS User Agent and MMS Center
- MM2 — the 3GPP interface between MMS Relay and MMS Server
- MM3 — the 3GPP interface between MMS Center and external servers
- MM4 — the 3GPP interface between MMS Centers
- MM5 — the 3GPP interface between MMS Center and HLR
- MM6 — the 3GPP interface between MMS Center and user databases
- MM7 — the 3GPP interface between MMS VAS applications and MMS Center
- MM8 — the 3GPP interface between MMS Center and the billing systems

Mobile Telephony – Short Codes

What Is It?

Short codes, also known short numbers or Common Short Codes (CSC) are special telephone numbers, significantly shorter than full telephone numbers, which can also be used to address SMS and MMS messages from mobiles or fixed phones.

They are designed to be shorter to read out and easier to remember than a normal length telephone number. While similar to telephone numbers, they are, at the technological level, unique to each operator, although providers generally have agreements to avoid overlaps.

Short codes are widely used for value-added services such as television voting, ordering ringtones, charity donations and mobile services (such as Google's SMS search service and 82ASK). Messages sent to short code numbers are generally billed at a higher rate than a standard SMS.

Regional Differences

Australia

Short Codes are 6 or 8 digits in length, both 6 and 8 digit shortcodes start with the pre-fix 19 followed be an additional 4 or 6 digits. Short Codes in Australia are commonly referred to as Premium Rate SMS Messages.

Canada

Codes are 5 or 6 digits in length. Codes starting with 4 are not permitted due to handset incompatibilities.

Czech Republic

Messages sent to/from these short codes are known as Premium Rate SMS. Codes are 7 digits in length for MO and 5 (not billed) or 8 (billed) for MT direction, starting with 9, while two or three (depending on billing type=MO/MT) trailing digits express the price, e.g. sms sent to 9090930 is billed for 30 CZK. Leading 3 digits are purpose type prefixes (908=micropayments, 909=adult content, 900=all other), digits at position 4 and 5 determinates the service provider registered by a network operator. There are also other 4digit shortcodes, used by a network operators for service only purposes (operator dependent)

Denmark

Codes are 4 digits in length.

Germany

Codes are 5 digits in length.

India

There are many players in the Indian Market who sells Keywords on these Shortcodes on Monthly Rental basis, Some of major big players are citisms.com, yahoo.co.in, hungama.com, rahuketu.com, Short Codes are 4 digits in length. The 4 digits can be further added with 4 more digits like wxyz, These Short Codes are Commonly referred to as Premium Rate SMS Messages and Costs around 80 paise to 6 Rs. Per SMS,

In late 2006, the regulator has asked all operators to switch to 5 digit Short Codes.

Indonesia

Codes are 4 digits in length with Rp. 2000,- (±$0.20) premium price.

New Zealand

Codes are 3 and 4 digits in length.

South Africa

Codes are 5 digits in length.

The Netherlands

Codes are 4 digits in length.

United Kingdom

Codes are 5 digits in length, starting with 6 or 8 (codes starting in 5 and 7 exist, but are reserved for future expansion). Individual network operators may come to an agreement with customers, allowing any number to be used - except for adult content services, which must use codes starting with 69 or 89.

United States Of America

As of May 31 2006 the standard lengths for interoperable Short Codes are 5 and 6 digits. Carriers use short codes with fewer digits for carrier specific programs - e.g., text 611 to see how many minutes you have remaining on your plan. Codes starting with 1 are not permitted.

Mobile Telephony – Mobile Marketing

What Is It?

Mobile Marketing can refer to one of two categories of marketing. First, and relatively new, is meant to describe marketing on or with a mobile device, such as a mobile phone. Second, and a more traditional definition, is meant to describe marketing in a moving fashion - for example - technology road shows or moving billboards.

Marketing on a mobile phone has become increasingly popular ever since the rise of SMS (Short Message Service) in the early 2000s in Europe and some parts of Asia when businesses started to collect mobile phone numbers and send off wanted (or unwanted) content.

Over the past few years SMS has become a legitimate advertising channel. This is due to the fact that unlike email over the public internet, the carrier who police their own networks have set guidelines and best practices for the mobile media industry (including mobile advertising). The IAB (Interactive Advertising Bureau) and the MMA (Mobile Marketing Association), as well, has established guidelines and evangelising the use of the mobile channel for marketers.

Mobile Marketing via SMS has expanded rapidly in Europe and Asia as a new channel to reach the consumer. SMS initially received negative media coverage in many parts of Europe for being a new form of spam as some advertisers purchased lists and sent unsolicited content to consumer's phones; however, as guidelines are put in place by the mobile operators, SMS has become the most popular branch of the Mobile Marketing industry with several 100 million advertising SMS sent out every month in Europe alone.

In North America the first cross-carrier SMS shortcode campaign was run by Labatt Brewing Company in 2002. Over the past few years mobile short codes have been increasingly popular as a new channel to communicate to the mobile consumer. Brands have begun to treat the mobile shortcode as a mobile domain name allowing the consumer to text message the brand at an event, in store and off any traditional media.

SMS services typically run off a short code, but sending text messages to an email address is another methodology. Short codes are 5 or 6 digit numbers that have been assigned by all the mobile operators in a given country for the use of brand campaign and other consumer services. The mobile operators vet every application before provisioning and monitor the service to make sure it does not diverge from its original service description.

One key criterion for provisioning is that the consumer opts in to the service. The mobile operators demand a double opt in from the consumer and the ability for the consumer to opt out of the service at any time by sending the word STOP via SMS. These guidelines are established in the MMA Consumer Best Practices Guidelines which are followed by all mobile marketers in the United States. The guidelines can be accessed at www.mmaglobal.com

Mobile Marketing Via MMS

Brands are delivering promotional content such as mobile music to mobile games to drive consumer engagement. This mobile content is delivered via MMS (Multimedia Message Service). Brands are also leveraging consumer-generated content.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

A good example of this is Motorola's ongoing campaigns at House of Blues venues where the brand allows the consumer to send their mobile photos to the LED board in real-time as well as blog their images online.

Mobile Marketing Via Bluetooth

The rise of Bluetooth started around 2003 and a few companies in Europe have started establishing successful businesses. Most of these businesses offer "Hotspot-Systems" which consist of some kind of content-management system with a Bluetooth distribution function. This technology has the advantages that it is permission-based, has higher transfer speeds and is also a radio-based technology and can therefore not be billed (i.e. is free of charge).

Mobile Marketing Via Infrared

Infrared is the oldest and most limited form of Mobile Marketing. Some European companies have experimented with "shopping window marketing" via free Infrared waves in the late 90s. However, Infrared has a very limited range (~ approx. 10cm - 1meter) and could never really establish itself as a leading Mobile Marketing technology.

User Controlled Media

Mobile marketing differs from most other forms of marketing communication in that it is often user (consumer) initiated, called Mobile Originated (or MO) message, and requires the express consent of the consumer to receive future communications. A call delivered from a server (business) to a user (consumer) is similarly called a Mobile Terminated (or MT) message. This infrastructure points to a trend set by mobile marketing of consumer controlled marketing communications. See also Push-Pull strategy on the nature of mobile marketing in practice by business.

Mobile Telephony – Mobile Learning

What Is It?

Mobile Learning (a.k.a. 'M-learning' or 'mLearning') is the follow up of E-learning which for its part originates from D-learning (distance education). M-learning is the delivery of learning to students who are not keeping a fixed location or through the use of mobile or portable technology. The rapid growth of information and communication technologies (ICT) makes it possible to develop new forms of this education. (Georgiev, T., E.Georgieva, A.Smrikarov. M-Learning - A New Stage of E-Learning, http://ecet.ecs.ru.acad.bg/cst04/Docs/sIV/428.pdf)

Definition

mLearning describes an array of ways that people learn, test or stay connected with knowledge needed for school, work or task accomplishment. Devices utilized include: Mobile Phones, PDAs (such as a Palm or Pocket PC) -- or the combination of the two in a Smart Phone -- and digital audio players such as an mp3 player. This can redefine "on the job" training for someone who accesses a lesson literally "just in time" while faced with a new challenge and they have to turn to their mobile device for instant answers. Field research can take a wealth of knowledge and data into the field with these devices and also send back new data instantly for others to use elsewhere. This is a form of eLearning where mobility matters and the connectedness while wandering away from a desktop or laptop plugged into a wired connection extends the usefulness and timeliness of the lesson and learning experience -- perhaps shared with other mobile learners.

Differences Between M-Learning & E-Learning

If e-learning took learning away from the classroom or campus, then m-learning is taking learning away from a fixed point. Where e-learning is an alternative to classroom learning (actually eLearning should/can be complementary to classroom learning) - m-learning is a complementary activity to both e-learning and traditional learning. M-learning respects that a user would like to interact with educational resources whilst away from their normal place of learning - classroom or computer.

In one sense m-learning has been around for longer than e-learning, with the paperback book and other portable resources, but technology is what shapes today's usage of m-learning. Technology now allows us to carry vast resources in our pockets and access these wherever we find convenient. Technology also allows us to interact with our peers instantaneously and work together remotely in ways never before possible.

While the opportunities that m-learning devices present us with are new - the challenges are quite old, smaller screen sizes, limited processing power, reduced input capabilities. These challenges mean that adapting existing e-learning services and content to m-learning is not a trivial task.

m-Learning has been gaining appeal among younger generations who have grown up using portable video game devices and wireless technology. In this sense, mLearning appeals not only to those who need learning portable, but to those who have grown up with a cognitive disposition towards using mobile devices- whether or not they have the need for true portability in their learning.

Challenges With M-Learning

The connectivity is one of the main differences if we compare a mobile device with the PC (the usual medium for delivering e-learning). Nowadays mobile devices might be connected to ‘The Net’ via many technologies – WAP, GPRS, UMTS, Bluetooth, WiFi, etc. Although it is predictable that in the future the ‘always on’ will be wide spread still it is not the case. Mobile devices often have periods of disconnection, either intentionally (when the connection is too expensive) or not (when no infrastructure is provided).

Devices’ hardware and software characteristics have a big impact on what content is possible and meaningful to be delivered. Usually the web content is designed for desktop PCs, thus unpleasant and even rarely useful from a small-screened device. Nowadays mobile phones are rapidly becoming increasingly powerful (both from hardware and software point of view) however their screens will remain comparatively small. Often also the navigation is hard. Equipped with a small phone-style keyboard or a touch-screen (for the PDAs) the users might lose more time in searching where on the page the information they need is than in reading it. We can imagine alternative ways of navigation, for example voice commands. The memory available on a mobile device is also relatively small. It is possible to use extension packs on some devices like PDAs, which reduces some of the restrictions.

Devices such as the Nintendo DS raise the bar even higher by providing collaborative gaming. The Nokia N-Gage brings an expandable platform of gaming and mobile communications to the forefront.

mp3 players and other portable media devices have also come into the forefront of the mLearning trend with companies using these devices to deploy education to their 'road warrior' sales teams through podcasts or videocasts.

Location is a new thing to be considered. Although up to now we are talking only about limitations confronting m-learning and e-learning, there are also advantages. The small size of the device and the wireless connections make them available anytime and anywhere. The mobility opens variety of new scenarios. Services involving location-discovery are, for example, receiving directions on how to get to a certain room, or alerts for seminars/lectures that can be triggered while taking into consideration the current place and the time to get to the needed destination; location-aware printing of the learning content, etc.

Source: Copied rather too directly from Trifonova A., Ronchetti M. (2004). "A General Architecture to Support Mobility in Learning". Proc. of ICALT 2004 [IEEE Computer Society Press 2004, ISBN 0-7695-2181-9]. pp. 26-30.

Different Communication Methods

- Text Message / SMS
- WAP Wireless Application Protocol
- GPRS
- Bluetooth
- WLAN / Wifi
- IrDA
- I-Phone

References

- B Sanregret, Hot Lava Software, Mobile Learning: A Reality, Arlington, VA, SALT Conference, 2006.
- L Low & M O'Connell, Learner-Centric Design of Digital Mobile Learning, Queensland University of Technology, 2006.
- Sharma, S., F.Kitchens. Web Services Architecture for M-Learning, Electronic Journal on e-Learning, Vol.2, Issue 1, 2004.
- T Georgiev, E Georgieva, A Smrikarov - Proceedings of the 5th international conference on Computer Systems and Technologies - CompSysTech'2004. http://ecet.ecs.ru.acad.bg/cst04/Docs/sIV/428.pdf
- A Trifonova, M Ronchetti. A General Architecture to Support Mobility in Learning. Proc. of the 4th IEEE ICALT 2004
- B Sanregret, Hot Lava Software, MLearning Basics 2005
- T Toth. Technology for Trainers, ASTD Press 2003
- D. McPhee and N. Thomas, "Evaluating the effectiveness of m-Learning in the teaching of multi-media to first year university students”, iJET International Journal of Emerging Technologies in Learning" Vol 1. 2006

Mobile Telephony – Mobile News

What Is It?

Mobile News services are growing in popularity along with an explosion in the usage of SMS messages worldwide and a few organizations are exploring these services. International news agency CNN advertises its mobile news services over its news broadcasts regularly.

Reuters also provides mobile alerts free to customers in the U.S. and Canada. Other institutions offering SMS news services include Channel News Asia Sky News Starcomm.

A small news agency Jasmine News in Sri Lanka is exploring public journalism using mobile telephony. Its only a matter of time before a number of news organizations explore public journalism via the medium of mobile phones.

Setting up SMS news services cost significantly less compared to setting up larger news organizations. Mobile News also has the potential to place the power of breaking news reporting in the hands of small communities and facilitate a much better exchange of information among users due to the ease of usage of mobile phones compared with conventional media such as radio, TV or newspapers.

SMS news services can be distinguished between "instant" and "on-demand" SMS news services, with instant pushed SMS messages having the most potential to transform public engagement with news.

Other institutions are exploring activism and the power for mobiles to record news in the hands of the public. Witness.org tries to empower people all over the world to tell their stories and share their experiences so that they have a voice.

Reuters and Yahoo have started You Witness News where news photos can be sent to Reuters at pics@reuters.com.

The Mobile Data Association (MDA) has forecast that SMS will become even more popular in 2007, reaching 45 billion for the year (3.75 billion a month, or 123 million a day).

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
*www.*planetlimited.*io*

Mobile Telephony – Mobile Publishing

What Is It?

Mobile publishing is the act of making something public through mobile phones.

One advantage for publishers is the ease of adding the cost of downloading content to the users' phone bills. Mobile music revenues (mostly from ringtones) amounted to $400m in 2005.

The arrival of more powerful telecoms networks such as 3G has enabled video publishing to mobile phones using tools such as FORscene. Video can be downloaded over the air, and once downloaded can be distributed virally over the phones' Bluetooth connections.

Examples

- Text: mainstream broadcasters publish news and other reports on mobile phones by text messaging.
- Ringtones: many operators provide ring tones for a fee.
- Wallpapers: operators provide wallpapers for a fee. But there are also many sites where they can be received for free.
- Video: organisations can now publish promotional content on mobile phones.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

Mobile Telephony – MoBlog

What Is It?

Moblog is a blend of the words mobile and weblog. A mobile weblog, or Moblog, consists of content posted to the Internet from a mobile or portable device, such as a cellular phone or PDA. Moblogs generally involve technology which allows publishing from a mobile device.

Much of the earliest development of Moblogs occurred in Japan, among the first countries in the world where camera phones (portable phones with built-in cameras) were widely commercially available.

According to Joi Ito's History of Moblogs, the first post to the web from a mobile user was from Steve Mann in 1995. He used a wearable computer, a more elaborate predecessor to modern Moblogging devices. The first post to the Internet from an ordinary mobile device is believed to be by Tom Vilmer Paamand in Denmark in May 2000.

The term "Moblogging" itself was coined by Adam Greenfield to describe the practice in 2002. He went on to organise the First International Moblogging Conference (or 1IMC) in July 2003 in Tokyo, with the help of Paul Baron, Gen Kanai, Carsten Schwesig and Steve Graff. Less known about is the First International Love Hotel Moblogging Conference that took place a day before the real 1IMC event.

The term is sometimes pronounced with the emphasis on the first syllable - MOBlog - out of affinity with the ideas about social self-organization developed in Howard Rheingold's "Smart Mobs".

Weblogs made from portable devices are also sometimes known as cyborgLogs, abbreviated as 'glogs, especially when primarily image-based.

Illicit and taboo activities have proven popularity in some early Moblog experiments, while family-oriented moblogs may be soon implementing advanced filter controls. See Drewing for more information about the delinquency publishing fad.

In 2004, Singapore launched a National Day Moblog on National Day. Apparently, it is the first national Moblog in the world.

Early on in Moblogging users sent their media to a Moblog server via MMS or email. Recently software has become available which allows people to have the same rich experience they had while blogging from their PC. Some countries are even using moblogs for pedagogical purposes. The Singapore Government oraganizes annual Campus Moblogging competitions between its primary and secondary schools www.campusmoblog.com.sg

The art of the Moblog is that "a picture tells a thousand word." By posting pictures the Moblogger is able to allow the viewer to look through their eyes, to visually experience where he or she is and what he or she is doing. Words often do not describe what a picture can do very easily. Moblogging is becoming more widespread through the use of websites where anybody with a cameraphone and the ability to send pictures can create an account and participate by uploading pictures on the fly.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

MoVlogging

MoVlog is the short form of mobile video blog. It is basically a combination of the terms moblog and vlog.

The term 'movlog' can be used either as a noun or as a verb. As a noun, it refers to the actual blog that contains video entries. As a verb, it refers to the act of video blogging using a mobile device (e.g., cellphones and PDAs).

Mobile Telephony – Mobile 2.0

What Is It?

Mobile 2.0, refers to the perceived next generation of mobile internet services including Social networking sites, wikis , that emphasise collaboration and sharing amongst users. In essence it refers to bringing Web 2.0 services to the mobile internet.

Characteristics Of Mobile 2.0

- Social networking on the mobile
- "Network as platform" — delivering (and allowing users to use) applications entirely through a mobile browser
- Extensive use of User-Generated Content, so that the site is owned by its contributors

Implementations Of Mobile 2.0

Mobile 2.0 is still at the discussion stage. One of the early innovators in this area is Red Circle Technologies. The company provides a mobile labs framework to support users developing Mobile 2.0 sites called Mobile Labs.

Understanding Mobile 2.0

Written by Rudy De Waele of m-trends.org and edited by Richard MacManus.

On the eve of Le Web 3 in Paris - and one month after the Web 2.0 Summit concluded - it seems like an appropriate time to explore the world of the mobile Web, a.k.a. mobile 2.0. There has been a lot of discussion lately on this topic, a good deal of it inspired by the mobile 2.0 event - a one-day event held on 6 November 2006, organized by Daniel Appelquist and Mike Rowehl.

Carriers & Mobile Operators Are Taking Notice...

In the closing session about carriers and operators at the Under The Radar: Mobility Conference on 16 November 2006, I heard an Executive Director from Verizon Wireless using the term "Mobile 2.0". Also Orange (France Telecom) is sponsoring one of the biggest web 2.0 related conferences in Europe, Le Web 3 in Paris. The fact that carriers/operators are now linking their brand name to web 2.0/mobile 2.0 related content and conferences, shows that progress is being made. Web 2.0 inspired projects going mobile and/or mobile 2.0 projects have been considered as things to avoid for carriers/operators up till now, since they are disruptive to their current business models.

So does this mean, with the carriers/operators entering the space now, that mobile 2.0 is finally taking off?

Definitely in Europe. What the Web 2.0 Summit completely ignored is being picked up by Le Web 3 conference organizer Loïc Le Meur, who is including a panel on Mobility 2.0. It's being run by Charlie Schick from Nokia and Marko Ahtisaari from Blyk (the 1st pan-European free mobile operator). Another panel features Jyri Engeström from Jaiku.com and Felix Petersen from Plazes.

What Is Mobile 2.0?

It's absolutely necessary that more connections are made between the players in the web 2.0 sphere (a.k.a. next generation web apps & services) and what some Mobilists are calling mobile 2.0. What we mean by 'mobile 2.0' is another (r)evolution, already started, that will dramatically change the web and the mobility landscape that we currently know. The idea is that the mobile web will become the dominant access method in many countries of the world, with devices that become more hybrid and networks that become more powerful - everywhere in the next decade to come.

The rapid penetration of Wireless Broadband Access (WBA) technologies such as 3G/UMTS, the migration of traditional telecom networks to internet technology, the availability of affordable and functional Wi-Fi and dual mode Wi-Fi/mobile phones... will all boost VoIP over broadband internet and ultimately blur the distinction between fixed and mobile services, since both become wireless and IP based.

I often ask myself the question of whether it'll be easier for web 2.0 apps to go mobile, or easier to create a mobile-specific web app or a service that can be easily connected with a web service? The answer is of course that both have a good chance to become even more important aspects of tomorrow's Web than they are now. Why? Well Eric Schmidt, CEO of Google, put it very well in a Financial Times article in May this year:

> “Mobile phones are cheaper than PCs, there are three times more of them, growing at twice the speed, and they increasingly have Internet access. What is more, the World Bank estimates that more than two-thirds of the world's population lives within range of a mobile phone network. Mobile is going to be the next big Internet phenomenon. It holds the key to greater access for everyone - with all the benefits that entails."

Obstacles To Overcome

Increasingly we assume that our PC is always connected, however the mobile device cannot yet guarantee such 'always on' connectivity - because the mobile network doesn't work the same way. This might be one of the few hurdles left to overcome for mobile 2.0 apps and services going mass market.

Non-carrier projects like Google Wi-Fi and FON aim to make cities completely Wi-Fi accessible. From personal experience I can tell you that people are going to use these alternative options to connect to the internet, once it's available on their mobile devices.

Mobilist blogger Enrique C. Ortiz sees another hindrance (and I think he's right): the lack of open standards and tools to build your own mobile 2.0 applications. He says:

> "Web 2.0 is based on user intelligence instead of technologies, i.e. by giving users smart tools that enable them to apply human semantics to information provided, you get a more intelligent web. This can only be done in a massive (thus useful) way with open standards and protocols that are inclusive and inviting to everyone. Now, as I see it, this 'open-source' story is an aspect seriously lacking from mobile platforms."

Carriers/operators need to cover their investments and so they want to be compensated by any 3rd party using their network. This is fair enough, but the fact is that operators are losing more and more control over mobile devices - because these devices can communicate with other devices over Bluetooth, Wi-Fi, Wimax, NFC, etc. That is, more options are becoming available for mobile users to access the web over networks other than the closed networks of the operators.

Internet Players Entering The Market

Another important thing happening is that handset manufacturers like Nokia, Motorola and Sony-Ericsson (to name the most important) are getting the company of new players like Apple (iPod), Microsoft (Zune) and other device companies, which are entering the mobility space and connecting the physical and the virtual worlds through the mobile. Not forgetting Google's possible strategy of offering free phones.

If you have had the opportunity to experience the latest Nokia NSeries phones, you have gotten close to understanding what mobile 2.0 is all about. It's about connecting your phone through Wi-Fi networks to browse the latest innovative, mobile accessible web 2.0 services. For example downloading your favourite podcasts, reading your RSS feeds, doing a one-click image upload to Flickr (nicely tagged with ShoZu), consulting the location map while on the road, tagging your streamed video's, etc.

Mobile Startups

There's definitely a lot of movement around on the mobile start-up front. Besides Yahoo with Flickr and Google with YouTube going mobile, there are some very interesting start-up companies resolutely going mobile. Many of them are building easy-to-use mobile web apps and services. Here's a starter to check out (too many to link to, but just google them!):

- BluePulse
- ComVu
- Funambol
- Gizmo
- Loopt
- JuiceCaster
- Mobo
- Mystrands
- Plazes
- Plusmo
- Sharpcast
- SlingMedia
- Shozu
- SoonR
- TalkPlus
- Widsets
- Winksite

Relationship Between Mobile 2.0 & Web 2.0

I'm not sure who coined the term 'mobile 2.0' first, but loads of discussions and conversations have been going on for a while now (see links below). To me, the shift happened at a Mobile Music event in London in November 2005. I presented a couple of slides trying to explain what I thought was happening at that point: a Fixed Internet Mobile and Network Convergence, combined with the coming of Hybrid Phones (Nokia no longer calls them phones!), combined with ability to access web 2.0 services.

At that point Tim O'Reilly had just released his article defining web 2.0 and Ajit Jaokar was about to write his first definitions on mobile 2.0, which later resulted in his book Mobile Web 2.0 - which explored the more in-depth relations between web 2.0 and mobile 2.0 related apps and services. His blog and book are both recommended reading for anyone interested in this topic. [Ed: later this week we'll feature an extract from Ajit's book here on Read/WriteWeb]

Here are some essential components of what mobile 2.0 is about:

1) Openness: open standards, open-source development and open access - creating more options for the user, not enclosing them in the walled gardens currently (still) used by operators.

2) The context of accessing the network and associated web services needs to be a positive user experience. For example for mobile search, the context includes: browser type, different device functionalities, security issues, display on a small screen, how to insert ads, etc. Associated with this is the usability experience of the devices, applications and services - and other components. For a more detailed analysis of context and the mobile web, see this article I wrote for gotomobile.

3) Affordable pricing to use the network to access content and services.

4) More user choice in the ways to communicate and share experiences with others (social interaction)

5) Intelligent 'aware' applications and devices that know where you are; location 'aware' applications seamlessly integrated.

6) New business opportunities coming to market, which may or may not connect to operators networks; think RSS feeds, alerts to SMS, Bluetooth and Wi-Fi Entertainment download zones and access spots, Podcasting to your mobile, Streaming Videocall to TV, Moblogging, Video blogging and media sharing applications, Click to Call (a phone number tagged into a mobile web or WAP page), Mobile Search, and last but not least VoIP tools & services.

Also worth noting that one-click access to the (mobile) web is essential to deploy easy-to-access online services. In this area there has been a lot of movement, with companies proposing solutions using QR codes, image recognition and augmented reality applications in mobile.

Conclusion

To conclude, check this nice mobile 2.0 definition from Daniel Appelquist:

> "Mobile 2.0 is not "the Future." it is services that already exist all around us. These services are maturing at an amazing rate and what they are doing is effectively knitting together Web 2.0 with the mobile platform to create something new: a new class of services that leverage mobility but are as

> easy to use and ubiquitous as the Web is today. These services point the way forward for the mobile data industry."

It took the internet a couple of years after 1994 to reach its maturity on the technology side, not to forget the business side of things. I believe the time has come for another exciting period, the Mobile Web. Some carriers/operators are finally starting to act - how about you?

Mobile Telephony – Mobile Social Networking

What Is It?

Mobile social networking is one or more individuals, with similar interests or commonalities, conversing and connecting with one another using the mobile phone. Much like Internet social networking, mobile social networking occurs in virtual communities. A current trend for Internet social networking websites such as MySpace and Facebook is to turn mobile. Exclusively mobile-based mobile social networks, such as those operated by AirG MocoSpace and Prodigits . Advances in software technology have facilitated the existence of these mobile virtual communities. Industry wireless network technologies include SMS, WAP, Java, BREW and i-mode.

Mobile Community

In most mobile communities, mobile phone users can now create their own profiles, make friends, participate in chat rooms, create chat rooms, hold private conversations, share photos, and share blogs by using their mobile phone. Mobile communities are becoming more and more popular among young consumers.

175 Million Users By 2010

Mobile 2.0 - Mobile Social Communities: 175 Million users by 2010

Wireless News, March 7th, 2007

Wireless Federation Research on ”Mobile Social Communities” revealed that there are currently nearly 45 million members using it worldwide, a number that is expected to reach 175 million in 2012.

“With a huge rise witnessed in Online Social communities like MySpace, Facebook among others, there is now a paradigm shift in how people connect with each other using the electronic medium.” says Mohit Sehgal of the Wireless Federation. Increasingly there will be a move to the mobile given the increasing availability of bandwidth from the mobile operators and readiness of handsets available.

“Such mobile social communities will extend the reach of electronic social interaction to millions who don’t have regular or easy access to computers. The most significant increase may be seen in countries like India which have a far higher mobile penetration as compared to PCs”

The rapid pace of mobile social community growth means opportunities for new entrants hoping to join the established players such as SMS.ac, AirG, and Jumbuck that provide the technology and marketing behind leading mobile communities.

Opportunities to monetize mobile social communities fall into several main categories:

- Mobile operators profit from data usage that underpins all mobile community activities they carry, and in some cases from monthly subscription fees as well.

- Companies can sponsor special interest communities that relate directly to their brands or services.
- The self-profiling nature of these communities means that advertising can be targeted to specific niches with great accuracy. Many mobile communities also offer downloadable merchandise for sale — ring tones or images, for example.

“What would help drive these mobile social communities is for more mobile operators to sponsor them”.

Vodafone announcing the launch of MySpace on their network is a positive sign and will shape the future for such similar relationships.

Mobile Telephony – MoSoSo

What Is It?

MoSoSo, or mobile social software, is software - generally on a mobile phone or on a laptop computer - that facilitates social encounters, or mobile social networking by associating geographical location and time with one's own social network.

An emerging trend within the MoSoSo community is MoSoSo advertising. There is already experimentation of this going on in the UK and Europe.

The Perfect MoSoMo

Posted by Russell Buckley on MobHabby, May 23, 2005

The MoSoSo (Mobile Social Software) arena is clearly hotting up right now, even more so following Google's purchase of Dodgeball a few weeks back. It's also interesting that a billion dollar entity like Google still buys tiny little two-men-in-a-garage startups.

So I thought I'd put down some of my thoughts on the sector.

Online YAFRO's (a term coined by VC's in response to the endless business plans submitted - Yet Another Friendster Rip Off) were the big boom and fizzle sector of last year. Sites like Friendster and Linked-In attracted significant funding without having much of a clue about a business model. This isn't such a bad thing - it's easy to forget that Hotmail, Google and Yahoo! before them all didn't have much of a business model on launch, which clearly hasn't held them back.

The problem is that online social networking still doesn't have much of a model. Linked In is trying to make money from the job search market, while the others play around with various subscription-based revenue generation ideas.

Can the MoSoSo concept fare any better?

The basic idea of a MoSoSo is to overlay a location and time element to the idea of digital networking. So it enables you to find people in your vicinity and at that time for social, sexual/ dating or business networking. It's worth noting that the time variable is often overlooked in analysis of MoSoSo dynamics.

Examples of MoSoSo usage might just be to find out which of your buddies are in the area you're drinking in. Or find friends of friends, on the basis that if you're both pals of the same person, you're likely to have something in common and unlikely that the other will rape you. Or you may want to find other people interested in collecting rare stamps, so you can say "Hi, lovely set of Penny Reds, you're got there!". Or find business contacts who are in the market for your services.

All this you can do with your mobile phone. You can use the low-tech Dodgeball way - essentially sending an SMS saying "I'm here, who else is nearby?". Or more sophisticated methods relying on location identification or the new Nokia Sensor - a Bluetooth-based solution.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

The main issue all these ideas face is the old "critical mass" chestnut. If you're out on the town and fancy a hot date or a cold beer with a pal, you might try this kind of service. But if it fails to turn up anyone, you're going to lose interest pretty quickly and forget to use it.

This isn't such a problem if the service is completely automated, simply requiring the use to react. This means that even if critical mass takes time to build, at some point the service will alert the user and stand a chance to gaining usage.

This element is also important anyway, as the best services work with, rather than against, inertia. If you can avoid making the user have to do anything (other than sign up or download an application) you're making your life and theirs much, much easier.

So I'd say the first requirement for a successful MoSoSo is an automated process, where upon joining or downloading the application, the phone, or server, does all the work.

The second really-nice-to-have would be an existing community to recruit from - the obvious one being an existing online social networking site. But others could include an online dating site or media owner - The FT would be a good one for business networking, for example.

The third element would be distance. Currently Bluetooth works over about 10m, which is far too short. When you factor in the time dimension (don't forget this is about people being in the same place simultaneously) you're simply never going to find a target to hit, unless the parameters are really, really wide - like "humanoid, male".

The exception to this might be at specialist networking events, like trade exhibitions, where a niche market may exist for this type of product.

So, it has to work over greater distances - maybe a square kilometer (?), certainly while critical mass is being built up.

Profiling is also key as the user needs to feel that she's in control and can change the parameters whenever she wishes.

The final area is the business model. The Nokia angle is to sell more handsets and maybe for the initial download of the application, although I doubt it, in this case. But the real gold is being able to charge by usage, rather than a one-off charge.

I think the answer lies in a hybrid of Premium SMS and phone based application. In other words, you're allowed top line information for free (eg that there are three available, blonde, female, Aquarians within 100 metres) but you have to pay a small fee to find out who they are and get contact details.

This area also might be ripe for sponsorship by the right brand, if its keen on being seen as a bringing people together or a business networking catalyst.

What have I missed out? Has anyone seen a MoSoSo that they think has legs?

Mobile Telephony – GPS2SMS

What Is It?

The term GPS2SMS is used to indicate the transfer of Gps-coordinates by means of SMS (text messages).

For this the following applications exists:

- burglar system for vehicles and vessels (alarm goes off if it is transported illegally)
- tracking & tracing (following an object, e.g. a sea container).
- geofencing (alarm goes off if someone enters a prohibited area).
- digital bread crumb track of a route followed.
- a message for a tow truck or emergency services (911 / 112) about the localization of broken down vehicle or a driver with problems.

SMS messages can be sent by several devices, as long as they have a GSM modem and a GPS receiver. Handheld devices exist, as well as built-in systems for vehicles and programmes for desktop PC, laptop, PDA and Smart phone. Some hobbyists build their own solutions, such as a small device that links a GSM phone to a GPS receiver (using a bluetooth, Wi-Fi or USB port).

For theft detection of a vehicle a device can be used that sends an alarm text message containing the coordinates of the object to be monitored, as soon as it leaves a certain protected area.

MMS can also be sent, using Exchangeable image file format (Exif).

More About GPS In General

The Global Positioning System (GPS) is currently the only fully functional Global Navigation Satellite System (GNSS). Utilizing a constellation of at least 24 medium Earth orbit satellites that transmit precise microwave signals, the system enables a GPS receiver to determine its location, speed and direction.

Developed by the United States Department of Defense, it is officially named NAVSTAR GPS (Contrary to popular belief, NAVSTAR is not an acronym for NAVigation Satellite Timing And Ranging, but simply a name given by Mr. John Walsh, a key decision maker when it came to the budget for the GPS program). The satellite constellation is managed by the United States Air Force 50th Space Wing. The cost of maintaining the system is approximately US$750 million per year, including the replacement of aging satellites, and research and development. Despite this fact, GPS is free for civilian use as a public good.

GPS has become a widely used aid to navigation worldwide, and a useful tool for map-making, land surveying, commerce, and scientific uses. GPS also provides a precise time reference used in many applications including scientific study of earthquakes, and synchronization of telecommunications networks.

Simplified Method Of Operation

A GPS receiver calculates its position by measuring the distance between itself and three or

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited.**io

more GPS satellites. Measuring the time delay between transmission and reception of each GPS microwave signal gives the distance to each satellite, since the signal travels at a known speed. The signals also carry information about the satellites' location. By determining the position of, and distance to, at least three satellites, the receiver can compute its position using trilateration. Receivers typically do not have perfectly accurate clocks and therefore track one or more additional satellites to correct the receiver's clock error.

System Segmentation

The current GPS consists of three major segments. These are the space segment (SS), a control segment (CS), and a user segment (US).

Space Segment

The space segment (SS) is composed of the orbiting GPS satellites, or Space Vehicles (SV) in GPS parlance. The GPS design calls for 24 SVs to be distributed equally among six circular orbital planes. The orbital planes are centered on the Earth, not rotating with respect to the distant stars. The six planes have approximately 55° inclination (tilt relative to Earth's equator) and are separated by 60° right ascension of the ascending node (angle along the equator from a reference point to the orbit's intersection).

Orbiting at an altitude of approximately 20,200 kilometers (12,600 miles or 10,900 nautical miles; orbital radius of 26,600 km (16,500 mi or 14,400 NM)), each SV makes two complete orbits each sidereal day, so it passes over the same location on Earth once each day. The orbits are arranged so that at least six satellites are always within line of sight from almost everywhere on Earth's surface.

As of April 2007, there are 30 actively broadcasting satellites in the GPS constellation. The additional satellites improve the precision of GPS receiver calculations by providing redundant measurements. With the increased number of satellites, the constellation was changed to a nonuniform arrangement. Such an arrangement was shown to improve reliability and availability of the system, relative to a uniform system, when multiple satellites fail.

Control Segment

The flight paths of the satellites are tracked by US Air Force monitoring stations in Hawaii, Kwajalein, Ascension Island, Diego Garcia, and Colorado Springs, Colorado, along with monitor stations operated by the National Geospatial-Intelligence Agency (NGA). The tracking information is sent to the Air Force Space Command's master control station at Schriever Air Force Base in Colorado Springs, which is operated by the 2d Space Operations Squadron (2 SOPS) of the United States Air Force (USAF). 2 SOPS contacts each GPS satellite regularly with a navigational update (using the ground antennas at Ascension Island, Diego Garcia, Kwajalein, and Colorado Springs). These updates synchronize the atomic clocks on board the satellites to within one microsecond and adjust the ephemeris of each satellite's internal orbital model. The updates are created by a Kalman Filter which uses inputs from the ground monitoring stations, space weather information, and various other inputs.

User Segment

The user's GPS receiver is the user segment (US) of the GPS system. In general, GPS receivers are composed of an antenna, tuned to the frequencies transmitted by the satellites, receiver-processors, and a highly-stable clock (often a crystal oscillator). They may also include a display for providing location and speed information to the user. A receiver is often described by its number of channels: this signifies how many satellites it can monitor simultaneously. Originally limited to four or five, this has progressively increased over the years so that, as of 2006, receivers typically have between twelve and twenty channels.

GPS receivers may include an input for differential corrections, using the RTCM SC-104 format. This is typically in the form of a RS-232 port at 4,800 bit/s speed. Data is actually sent at a much lower rate, which limits the accuracy of the signal sent using RTCM. Receivers with internal DGPS receivers can outperform those using external RTCM data. As of 2006, even low-cost units commonly include Wide Area Augmentation System (WAAS) receivers.

Many GPS receivers can relay position data to a PC or other device using the NMEA 0183 protocol. NMEA 2000 is a newer and less widely adopted protocol. Both are proprietary and controlled by the US-based National Marine Electronics Association. References to the NMEA protocols have been compiled from public records, allowing open source tools like gpsd to read the protocol without violating intellectual property laws. Other proprietary protocols exist as well, such as the SiRF and MTK protocols. Receivers can interface with other devices using methods including a serial connection, USB or Bluetooth.

Navigation Signals

Each GPS satellite continuously broadcasts a Navigation Message at 50 bps giving the time, an almanac and an ephemeris. The almanac consists of coarse orbit and status information for each satellite in the constellation; a complete almanac transmission takes 12.5 minutes, and is responsible for the long initial acquisition process when a new receiver is first turned on. The ephemeris gives the satellite's own precise orbit and is transmitted every 30 seconds. The almanac assists in the acquisition of other satellites, while an ephemeris from each satellite is needed to compute position fixes using that satellite. The ephemeris is updated every 2 hours and is valid for 4 hours. The time needed to acquire it is a significant element of the delay to first position fix when a receiver is switched on after having been off for several hours.

Each satellite transmits its navigation message with at least two distinct spread spectrum codes: the Coarse / Acquisition (C/A) code, which is freely available to the public, and the Precise (P) code, which is usually encrypted and reserved for military applications. The C/A code is a 1,023 chip pseudo-random (PRN) code at 1.023 million chips/sec so that it repeats every millisecond. Each satellite has its own C/A code so that it can be uniquely identified and received separately from the other satellites transmitting on the same frequency. The P-code is a 10.23 megachip/sec PRN code that repeats only every week. When the "anti-spoofing" mode is on, as it is in normal operation, the P code is encrypted by the Y-code to produce the P(Y) code, which can only be decrypted by units with a valid decryption key. Both the C/A and P(Y) codes impart the precise time-of-day to the user. Frequencies used by GPS include:

- L1 (1575.42 MHz): Mix of Navigation Message, coarse-acquisition (C/A) code and encrypted precision P(Y) code.
- L2 (1227.60 MHz): P(Y) code, plus the new L2C code on the Block IIR-M and newer satellites.
- L3 (1381.05 MHz): Used by the Defense Support Program to signal detection of missile launches, nuclear detonations, and other high-energy infrared events.
- L4 (1379.913 MHz): Being studied for additional ionospheric correction.
- L5 (1176.45 MHz): Proposed for use as a civilian safety-of-life (SoL) signal (see GPS modernization). This frequency falls into an internationally protected range for aeronautical navigation, promising little or no interference under all circumstances.

The first Block IIF satellite that would provide this signal is set to be launched in 2008.

Calculating Positions

The coordinates are calculated according to the World Geodetic System WGS84 coordinate system. To calculate its position, a receiver needs to know the precise time. The satellites are equipped with extremely accurate atomic clocks, and the receiver uses an internal crystal oscillator-based clock that is continually updated using the signals from the satellites.

The receiver identifies each satellite's signal by its distinct C/A code pattern, then measures the time delay for each satellite. To do this, the receiver produces an identical C/A sequence using the same seed number as the satellite. By lining up the two sequences, the receiver can measure the delay and calculate the distance to the satellite, called the pseudorange.

The orbital position data from the Navigation Message is then used to calculate the satellite's precise position. Knowing the position and the distance of a satellite indicates that the receiver is located somewhere on the surface of an imaginary sphere centered on that satellite and whose radius is the distance to it. When four satellites are measured simultaneously, the intersection of the four imaginary spheres reveals the location of the receiver. Receivers known to be near sea level can substitute the sphere of the planet for one satellite by using their altitude. Often, these spheres will overlap slightly instead of meeting at one point, so the receiver will yield a mathematically most-probable position (and often indicate the uncertainty).

Calculating a position with the P(Y) signal is generally similar in concept, assuming one can decrypt it. The encryption is essentially a safety mechanism: if a signal can be successfully decrypted, it is reasonable to assume it is a real signal being sent by a GPS satellite. In comparison, civil receivers are highly vulnerable to spoofing since correctly formatted C/A signals can be generated using readily available signal generators. RAIM features do not protect against spoofing, since RAIM only checks the signals from a navigational perspective.

Accuracy & Error Sources

The position calculated by a GPS receiver requires the current time, the position of the satellite and the measured delay of the received signal. The position accuracy is primarily dependent on the satellite position and signal delay.

To measure the delay, the receiver compares the bit sequence received from the satellite with an internally generated version. By comparing the rising and trailing edges of the bit transitions, modern electronics can measure signal offset to within about 1% of a bit time, or approximately 10 nanoseconds for the C/A code. Since GPS signals propagate nearly at the speed of light, this represents an error of about 3 meters. This is the minimum error possible using only the GPS C/A signal.

Position accuracy can be improved by using the higher-speed P(Y) signal. Assuming the same 1% bit time accuracy, the high frequency P(Y) signal results in an accuracy of about 30 centimeters.

Electronics errors are one of several accuracy-degrading effects outlined in the table below. When taken together, autonomous civilian GPS horizontal position fixes are typically accurate to about 15 meters (50 ft). These effects also reduce the more precise P(Y) code's accuracy.

Sources Of User Equivalent Range Errors (UERE)

- * Ionospheric effects ± 5 meter
- * Ephemeris errors ± 2.5 meter
- * Satellite clock errors ± 2 meter
- * Multipath distortion ± 1 meter
- * Tropospheric effects ± 0.5 meter
- * Numerical errors ± 1 meter or less

Atmospheric Effects

Inconsistencies of atmospheric conditions affect the speed of the GPS signals as they pass through the Earth's atmosphere and ionosphere. Correcting these errors is a significant challenge to improving GPS position accuracy. These effects are smallest when the satellite is directly overhead and become greater for satellites nearer the horizon since the signal is affected for a longer time. Once the receiver's approximate location is known, a mathematical model can be used to estimate and compensate for these errors.

Because ionospheric delay affects the speed of microwave signals differently based on frequency—a characteristic known as dispersion—both frequency bands can be used to help reduce this error. Some military and expensive survey-grade civilian receivers compare the different delays in the L1 and L2 frequencies to measure atmospheric dispersion, and apply a more precise correction. This can be done in civilian receivers without decrypting the P(Y) signal carried on L2, by tracking the carrier wave instead of the modulated code. To facilitate this on lower cost receivers, a new civilian code signal on L2, called L2C, was added to the Block IIR-M satellites, which was first launched in 2005. It allows a direct comparison of the L1 and L2 signals using the coded signal instead of the carrier wave.

The effects of the ionosphere generally change slowly, and can be averaged over time. The effects for any particular geographical area can be easily calculated by comparing the GPS-measured position to a known surveyed location. This correction is also valid for other receivers in the same general location. Several systems send this information over radio or other links to allow L1 only receivers to make ionospheric corrections. The ionospheric data are transmitted via satellite in Satellite Based Augmentation Systems such as WAAS, which transmits it on the GPS frequency using a special pseudo-random number (PRN), so only one antenna and receiver are required.

Humidity also causes a variable delay, resulting in errors similar to ionospheric delay, but occurring in the troposphere. This effect is both more localized and changes more quickly than ionospheric effects and is not frequency dependent. These traits making precise measurement and compensation of humidity errors more difficult than ionospheric effects.

Changes in altitude also change the amount of delay due to the signal passing through less of the atmosphere at higher elevations. Since the GPS receiver computes its approximate altitude, this error is relatively simple to correct.

Multipath Effects

GPS signals can also be affected by multipath issues, where the radio signals reflect off surrounding terrain; buildings, canyon walls, hard ground, etc. These delayed signals can cause inaccuracy. A variety of techniques, most notably narrow correlator spacing, have been developed to mitigate multipath errors. For long delay multipath, the receiver itself can recognize the wayward signal and discard it. To address shorter delay multipath from the signal reflecting off the ground, specialized antennas may be used. Short delay reflections are harder to filter out since they are only slightly delayed, causing effects almost indistinguishable from routine fluctuations in atmospheric delay.

Multipath effects are much less severe in moving vehicles. When the GPS antenna is moving, the false solutions using reflected signals quickly fail to converge and only the direct signals result in stable solutions.

Ephemeris & Clock Errors

The navigation message from a satellite is sent out only every 12.5 minutes. In reality, the data contained in these messages tend to be "out of date" by an even larger amount. Consider the case when a GPS satellite is boosted back into a proper orbit; for some time following the maneuver, the receiver's calculation of the satellite's position will be incorrect until it receives another ephemeris update. The onboard clocks are extremely accurate, but they do suffer from some clock drift. This problem tends to be very small, but may add up to 2 meters (6 ft) of inaccuracy.

This class of error is more "stable" than ionospheric problems and tends to change over days or weeks rather than minutes. This makes correction fairly simple by sending out a more accurate almanac on a separate channel.

Selective Availability

The GPS includes a feature called Selective Availability (SA) that introduces intentional, slowly changing random errors of up to a hundred meters (300 ft) into the publicly available navigation signals to confound, for example, guiding long range missiles to precise targets. Additional accuracy was available in the signal, but in an encrypted form that was only available to the United States military, its allies and a few others, mostly government users.

SA typically added signal errors of up to about 10 meters (30 ft) horizontally and 30 meters (100 ft) vertically. The inaccuracy of the civilian signal was deliberately encoded so as not to change very quickly, for instance the entire eastern U.S. area might read 30 m off, but 30 m off everywhere and in the same direction. To improve the usefulness of GPS for civilian navigation, Differential GPS was used by many civilian GPS receivers to greatly improve accuracy.

During the Gulf War, the shortage of military GPS units and the wide availability of civilian ones among personnel resulted in a decision to disable Selective Availability. This was ironic, as SA had been introduced specifically for these situations, allowing friendly troops to use the signal for accurate navigation, while at the same time denying it to the enemy. But since SA was also denying the same accuracy to thousands of friendly troops, turning it off or setting it to an error of zero meters (effectively the same thing) presented a clear benefit.

In the 1990s, the FAA started pressuring the military to turn off SA permanently. This would save the FAA millions of dollars every year in maintenance of their own radio navigation systems. The military resisted for most of the 1990s, but SA was eventually "discontinued"; the amount of error added was "set to zero" at midnight on May 1, 2000 following an announcement by U.S. President Bill Clinton, allowing users access to the error-free L1 signal. Per the directive, the induced error of SA was changed to add no error to the public signals (C/A code). Selective Availability is still a system capability of GPS, and error could, in theory, be reintroduced at any time. In practice, in view of the hazards and costs this would induce for US and foreign shipping, it is unlikely to be reintroduced, and various government agencies, including the FAA, have stated that it is not intended to be reintroduced.

The US military has developed the ability to locally deny GPS (and other navigation services) to hostile forces in a specific area of crisis without affecting the rest of the world or its own military systems.

One interesting side effect of the Selective Availability hardware is the capability to correct the frequency of the GPS caesium and rubidium clocks to an accuracy of approximately 2 × 10-13 (one in five trillion). This represented a significant improvement over the raw accuracy of the clocks.

Relativity

According to the theory of relativity, due to their constant movement and height relative to the Earth-centered inertial reference frame, the clocks on the satellites are affected by their speed (special relativity) as well as their gravitational potential (general relativity). For the GPS satellites, general relativity predicts that the atomic clocks at GPS orbital altitudes will tick more rapidly, by about 45,900 nanoseconds (ns) per day, because they are in a weaker gravitational field than atomic clocks on Earth's surface. Special relativity predicts that atomic clocks moving at GPS orbital speeds will tick more slowly than stationary ground clocks by about 7,200 ns per day. When combined, the discrepancy is 38 microseconds per day; a difference of 4.465 parts in 1010. To account for this, the frequency standard onboard each satellite is given a rate offset prior to launch, making it run slightly slower than the desired frequency on Earth; specifically, at 10.22999999543 MHz instead of 10.23 MHz.

GPS observation processing must also compensate for another relativistic effect, the Sagnac effect. The GPS time scale is defined in an inertial system but observations are processed in an Earth-centered, Earth-fixed (co-rotating) system, a system in which simultaneity is not uniquely defined. The Lorentz transformation between the two systems modifies the signal run time, a correction having opposite algebraic signs for satellites in the Eastern and Western celestial hemispheres. Ignoring this effect will produce an east-west error on the order of hundreds of nanoseconds, or tens of meters in position.

The atomic clocks on board the GPS satellites are precisely tuned, making the system a practical engineering application of the scientific theory of relativity in a real-world system.

Techniques To Improve Accuracy

Augmentation

Augmentation methods of improving accuracy rely on external information being integrated into the calculation process. There are many such systems in place and they are generally named or described based on how the GPS sensor receives the information. Some systems transmit additional information about sources of error (such as clock drift, ephemeris, or ionospheric delay), others provide direct measurements of how much the signal was off in the past, while a third group provide additional navigational or vehicle information to be integrated in the calculation process.

Examples of augmentation systems include the Wide Area Augmentation System, Differential GPS, and Inertial Navigation Systems.

Precise Monitoring

The accuracy of a calculation can also be improved through precise monitoring and measuring of the existing GPS signals in additional or alternate ways.

The first is called Dual Frequency monitoring, and refers to systems that can compare two or more signals, such as the L1 frequency to the L2 frequency. Since these are two different frequencies, they are affected in different, yet predictable ways by the atmosphere and objects around the receiver. After monitoring these signals, it is possible to calculate and nullify that error.

Receivers that have the correct decryption key can relatively easily decode the P(Y)-code transmitted on both L1 and L2 to measure the error. Receivers that do not possess the key can still use a process called codeless to compare the encrypted information on L1 and L2 to gain much of the same error information. However, this technique is currently limited to specialized surveying equipment. In the future, additional civilian codes are expected to be transmitted on the L2 and L5 frequencies (see #GPS modernization, below). When these become operational, all users will be able to make the same comparison and directly measure some errors.

A second form of precise monitoring is called Carrier-Phase Enhancement (CPGPS). The error, which this corrects, arises because the pulse transition of the PRN is not instantaneous, and thus the correlation (satellite-receiver sequence matching) operation is imperfect. The CPGPS approach utilizes the L1 carrier wave, which has a period 1000 times smaller than that of the C/A bit period, to act as an additional clock signal and resolve the uncertainty. The phase difference error in the normal GPS amounts to between 2 and 3 meters (6 to 10 ft) of ambiguity. CPGPS working to within 1% of perfect transition reduces this error to 3 centimeters (1 inch) of ambiguity. By eliminating this source of error, CPGPS coupled with DGPS normally realizes between 20 and 30 centimeters (8 to 12 inches) of absolute accuracy.

Relative Kinematic Positioning (RKP) is another approach for a precise GPS-based positioning system. In this approach, determination of range signal can be resolved to an accuracy of less than 10 centimeters (4 in). This is done by resolving the number of cycles in which the signal is transmitted and received by the receiver. This can be accomplished by using a combination of differential GPS (DGPS) correction data, transmitting GPS signal phase information and ambiguity resolution techniques via statistical tests—possibly with processing in real-time (real-time kinematic positioning, RTK).

GPS Time & Date

While most clocks are synchronized to Coordinated Universal Time (UTC), the Atomic clocks on the satellites are set to GPS time. The difference is that GPS time is not corrected to match the rotation of the Earth, so it does not contain leap seconds or other corrections which are periodically added to UTC. GPS time was set to match Coordinated Universal Time (UTC) in 1980, but has since diverged. The lack of corrections means that GPS time remains at a constant offset (19 seconds) with International Atomic Time (TAI). Periodic corrections are performed on the on-board clocks to correct relativistic effects and keep them synchronized with ground clocks.

The GPS navigation message includes the difference between GPS time and UTC, which as of 2006 is 14 seconds. Receivers subtract this offset from GPS time to calculate UTC and specific timezone values. New GPS units may not show the correct UTC time until after receiving the UTC offset message. The GPS-UTC offset field can accommodate 255 leap seconds (eight bits) which, at the current rate of change of the Earth's rotation, is sufficient to last until the year 2330.

As opposed to the year, month, and day format of the Julian calendar, the GPS date is expressed as a week number and a day-of-week number. The week number is transmitted as a ten-bit field in the C/A and P(Y) navigation messages, and so it becomes zero again every 1,024 weeks (19.6 years). GPS week zero started at 00:00:00 UTC (00:00:19 TAI) on January 6, 1980 and the week number became zero again for the first time at 23:59:47 UTC on August 21, 1999 (00:00:19 TAI on August 22, 1999). To determine the current Gregorian date, a GPS receiver must be provided with the approximate date (to within 3,584 days) to correctly translate the GPS date signal. To address this concern the modernized GPS navigation messages use a 13-bit field, which only repeats every 8,192 weeks (157 years), and will not return to zero until near the year 2137.

Other Systems

Other satellite navigation systems in use or various states of development include:

- Beidou — China's regional system that China has proposed to expand into a global system.
- Galileo — a proposed global system being developed by the European Union, joined by China, Israel, India, Morocco, Saudi Arabia and South Korea, Ukraine planned to be operational by 2010.
- GLONASS — Russia's global system which is being restored to full availability in partnership with India.
- Indian Regional Navigational Satellite System (IRNSS) — India's proposed regional system.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

Mobile Telephony – Reverse SMS Billing

What Is It?

Reverse SMS billing means that the user of the recipient phone rather than the message sender is charged for the cost of the SMS message received. It is also called MT (Mobile Terminated) Billing. Reverse billed SMS messages are only sent if specifically requested by the phone user. A daily horoscope alert is an example of a service where regular reverse-billed messages are received.

Also called “premium sms service”.

In some countries, providers are required by law or regulation to provide a means of opting out of a service once it has been commenced. Most typically, this is achieved by sending a stop message (most typically, simply STOP) to the same number as the service itself.

Mobile Telephony – SMS Gateways

What Is It?

SMS gateways allow for the sending and receiving of SMS messages to or from devices other than cellphones.

Web Service To SMS

Aside from web-based gateways intended to be visited by end-users with web browsers, there are also web services that programs can utilize to send individual or bulk SMSes. Some companies offer web services with full APIs for managing SMS sending and accounting activities, and offer special options for transferring ringtones, etc.

SMS To Email

This type of service allows people who do not have a computer (or do not have current access to one), but do have a cellphone, to send email. Most U.S. carriers and recent phones have built-in support for sending email through such gateways. A service from the British company Connectotel, called M-Mail, allows for the sending of email messages even from phones and carriers that do not directly support this.

SMS To Landline & Landline To SMS

These services allow cellphone users to send SMS messages to landline phone numbers just as they would to other cellphones. With a representative service, Sprint's Text to Landline, after the customer has sent off the SMS message to the landline number, the recipient's phone rings with the caller ID of the Sprint customer's cellphone. When they pick up, an automated voice reads the text message and allows for a response via a voicemail or via one of a few canned text messages.

Several operators, including BT, Telefonica and Telecom Italia, have true fixed-wire SMS services. These are based on extensions to the ETSI GSM SMS standards and allow fixed-fixed, fixed-mobile and mobile-fixed messaging. These use Frequency-shift keying to transfer the message between the terminal and the SMSC. Terminals are usually DECT-based, but wired handsets and wired text-only (no voice) devices exist. Messages are received by the terminal recognising that the CLI is that of the SMSC and going off-hook silently to receive the message.

When messages are addressed to a device that lacks the ability to receive SMS, then a text-to-speech gateway is employed. The translated message is then either stored in the subscriber's voice mail-box, or the system places a call direct to the end-point and plays the message.

SMS To Skype

SMS to Skype, which is built on top of the M-Mail service mentioned above, allows any GSM mobile phone user to send Skype chat messages from their phone.

Skype To SMS

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www*.planetlimited.*io

The Skype client software, from version 2.5 onwards, provides access to the Skype SMS service, allowing Skype users to send SMS messages and charge their SkypeOut account.

SMS To Home Automation Systems

Some Home Automation systems have the ability to accept SMS messages from the users mobile phones to control lighting and other appliances. The system may send an SMS confirmation when the requested action has been performed

Free Third-Party Email To SMS Gateways

- FREE SMS (Free WorldWide SMS)
- Teleflip (North American carriers)
- Free Text Message (Free SMS to UK and Worldwide)
- PHP Text Message Open Source email to SMS

Free Third-Party Web To SMS Gateways

- Zemble.com (Free Text Messaging and Group Text Messaging to North America - No Need to Know the Recipient's Carrier)
- SEA SMS (Free SMS Worldwide, Registration NOT Required)
- Text 4 Free (Free WorldWide Text Messages, Plus Bulk Messaging - No Signups)
- Wadja FREE SMS (Mobile social network with Free SMS to Worldwide)
- SMS to RUSSIA (Free SMS to Russia and Ukraine)
- Free Text Messaging Service(Sending SMS to 190+ countries worldwide)
- Text for Free (Free worldwide text messages in 4 easy steps)
- Funtonia.com (Free SMS to North American, U.K., and Australian carriers)

Commercial Sservices

- ftp.2.sms - Software for bulk messaging with global coverage
- ViaNett - Resources and gateway for building sms-applications
- TextHQ - Send text and picture messages to mobile phones around the world from any computer
- SMStrend.net High quality Bulk SMS, web2SMS, email2SMS (SMS to North American, U.K, Italy, Germany and Spain)
- hoaxMail - Sending SMS from a user definable number
- Txtlocal SMS Marketing - Low cost international SMS, mail merge, support for 640 characters, free API, free setup
- TraiTel Telecommunications - Two way, international SMS operator for bulk sending. HTTP, SMPP

and SMTP API's at no extra charge.

- Human Mobile SMS - Gateway and bulk messaging for Brazilian carriers.
- Oksijen SMSC, SMPPGW - SMSC and SMPP GW for Network Operators
- MessageVista Communication gateway access API - HTTP API and web
- SMS Based Communication Software -
- Mobilant: Worldwide SMS Business Gateway

service gateways with international coverage.

- 160.com.au Web 2 SMS - Australia's cheapest web to SMS gateway offering excellent features, reliable network at the lowest message cost.
- Forest Interactive - Offering Premium SMS Service for Malaysia, Indonesia, Singapore, Thailand and 9 European countries. Provides interactive TV applications and also bulk sms services for a good rate.

Mobile Telephony – Short Message Service Centre

What Is It?

A Short Message Service Center (SMSC) is a network element in the mobile telephone network which delivers SMS messages.

Operation

When a user sends a text message (SMS message) to another user, the message gets stored in the SMSC which delivers it to the destination user when they are available. This is a store and forward operation. The SMSC usually has a configurable time limit for how long it will store the message, and users can usually specify a shorter time limit if they want.

A message may also come from an application, for example voice mail server sending voice mail incoming message alerts. Mobile operators allow businesses to interact with their SMSC to submit the messages in bulk. From SMSC point of view, such applications are called SME (Short Message Entities). In this case the SMSC is responsible for locating SMSC of the destination user and submitting the message there.

Protocols

For bulk transmission and reception of SMS messages, SMSCs have conventional, fixed, network interfaces as well as mobile network interfaces. A number of protocols have been defined to support this sort of wire-line access:

- SMPP (Short message peer-to-peer) - The most common protocol and the only non-proprietary one in wide use.
- EMI/UCP (External Machine Interface/Universal Computer Protocol) - A proprietary protocol by market leader LogicaCMG.
- CIMD (Computer Interface to Message Distribution) - A proprietary protocol developed by Nokia for its Artuse SMSC.
- OIS (Open Interface Specification) - Despite the name, a proprietary protocol developed by Sema Group (now Airwide Solutions).

Companies

Two companies who make many of the SMSCs in use in the GSM world are Logica and CMG - or rather just one company, LogicaCMG, as these two merged on January 1st, 2003. Other large players include Jataayu,Bharti Telesoft,Tele DNA (India) , Intervoice, Comverse Technology, Nokia, Huawei, Unisys, Airwide Solutions, Jinny, Oksijen Teknoloji, Telenity, Lucent and Motorola.

Other Companies working on SMSC development [Pyro SMSC] www.pyronetworks.com

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

Mobile Telephony – SMPP

What Is It?

The short message peer-to-peer protocol (SMPP) is a telecommunications industry protocol for exchanging SMS messages between SMS peer entities such as short message service centres. It is often used to allow third parties (e.g. value-added service providers like news organisations) to submit messages, often in bulk.

SMPP was originally designed by Aldiscon, a small Irish company that was later bought by Logica, now LogicaCMG. In 1999, SMPP was formally handed over to the SMPP Developers Forum, later renamed as The SMS Forum.

The protocol is based on pairs of request/response PDUs (protocol data units, or packets) exchanged over OSI layer 4 (TCP session or X.25 SVC3) connections. PDUs are binary encoded for efficiency.

The most commonly used versions of SMPP are v3.3, the most widely supported standard, and v3.4, which adds transceiver support (single connections that can send and receive messages). Data exchange may be synchronous, where each peer must wait for a response for each PDU being sent, and asynchronous, where receiving and transmitting go in independent threads with the use of buffers and timers. The latest version of SMPP is v5.0.

Example

This is an example of a 60-octet submit_sm PDU...

Hexdump

00 00 00 3C 00 00 00 04 00 00 00 00 00 00 00 05 00 02 08 35 35 35 00 01 01
35 35 35 35 35 35 35 35 35 00 00 00 00 00 00 00 00 00 00 0F 48 65 6C 6C 6F
20 77 69 6B 69 70 65 64 69 61

PDU Header

'command_length', (60) ... 00 00 00 3C
'command_id', (4) ... 00 00 00 04
'command_status', (0) ... 00 00 00 00
'sequence_number', (5) ... 00 00 00 05

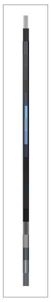
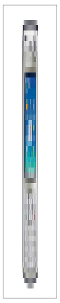
PDU Body

'service_type', () ... 00
'source_addr_ton', (2) ... 02
'source_addr_npi', (8) ... 08
'source_addr', (555) ... 35 35 35 00
'dest_addr_ton', (1) ... 01
'dest_addr_npi', (1) ... 01
'dest_addr', (555555555) ... 35 35 35 35 35 35 35 35 35 00
'esm_class', (0) ... 00
'protocol_id', (0) ... 00

'priority_flag', (0) ... 00
'schedule_delivery_time', () ... 00
'validity_period', () ... 00
'registered_delivery', (0) ... 00
'replace_if_present_flag', (0) ... 00
'data_coding', (0) ... 00
'sm_default_msg_id', (0) ... 00
'sm_length', (15) ... 0F
'short_message', (Hello wikipedia) ... 48 65 6C 6C 6F 20 77 69 6B 69 70 65 64 69 61'

Mobile Telephony – Mobile Phone

What Is It?

A mobile or cellular telephone is a long-range, portable electronic device for personal telecommunications over long distances. In addition to the standard voice function of a telephone, current mobile phones can support many additional services such as SMS for text messaging, email, packet switching for access to the Internet, and MMS for sending and receiving photos and video. Most current mobile phones connect to a cellular network of base stations (cell sites), which is in turn interconnected to the public switched telephone network (PSTN) (the exception are satellite phones).

Mobile phones are distinct from cordless telephones, which generally operate only within a limited range of a specific base station. Technically, the term mobile phone includes such devices as satellite phones and pre-cellular mobile phones such as those operating via MTS which do not have a cellular network, whereas the related term cell(ular) phone does not. In practice, the two terms are used nearly interchangeably.

History

The Mobile phone is one of the most used pieces of equipment today. The concept of using hexagonal cells for mobile phone base stations was invented in 1947 by Bell Labs engineers at AT&T (see History of mobile phones) and was further developed by Bell Labs during the 1960s. Radiophones have a long and varied history that stretches back to the 1950s, with hand-held cellular radio devices being available since 1983. Due to their low establishment costs and rapid deployment, mobile phone networks have since spread rapidly throughout the world, outstripping the growth of fixed telephony

In 1946, the 0G generation of mobile telephones were introduced. 0G mobile telephones, such as Mobile Telephone Service, were not officially categorized as mobile phones, since they did not support the automatic change of channel frequency in the middle of a call, when the user moved from one cell (base station coverage area) to another cell, a feature called "handover".

In 1970 Amos Joel of Bell Labs invented the "call handoff" feature, which allowed a mobile-phone user to travel through several cells during the same conversation. Martin Cooper of Motorola is widely considered to be the inventor of the first practical mobile phone for handheld use in a non-vehicle setting. Using a modern, if somewhat heavy portable handset, Cooper made the first call on a handheld mobile phone on April 3, 1973. At the time he made his call, Cooper was working as Motorola's General Manager of its Communications Division.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited*****.io***

Fully automatic cellular networks were first introduced in the early to mid-1980s (the 1G generation). The first fully automatic mobile phone system was the 1981 Nordic Mobile Telephone (NMT) system. Until the late 1980s, most mobile phones were too large to be carried in a jacket pocket, so they were usually permanently installed in vehicles as car phones. With the advance of miniaturization and smaller digital components, mobile phones got smaller and lighter.

Manufacturers

Nokia Corporation is currently the world's largest manufacturer of mobile telephones, with a global market share of approximately 36% in Q4 of 2006. Other mobile phone manufacturers include Apple Inc., Audiovox (now UT Starcom), Benefon, BenQ-Siemens, High Tech Computer Corporation, Fujitsu, Kyocera, 3G, LG, Motorola, NEC, HTC, Panasonic (Matsushita Electric), Pantech Curitel, Philips, Research in Motion, Sagem, Samsung, Sanyo, Sharp, Siemens, Sierra Wireless, SK Teletech, Sony Ericsson, T&A Alcatel, and Toshiba. There are also specialist communication systems related to, but distinct from mobile phones, such as for exampleProfessional Mobile Radio.

Subscriptions

In the UK, the number of cell numbers has surpassed the number of people. There will be over four hundred million cell phone users in China by 2008. Luxembourg has the highest mobile phone penetration rate in the world, at 164% in December 2005. In Hong Kong the penetration rate reached 117% of population in September 2004. The total number of mobile phone subscribers in the world was estimated at 2.14 billion in 2005. Around 80% of world's population have mobile phone coverage as of 2006. This figure is expected to increase to 90% by the year 2010.

At present, Africa has the largest growth rate of cellular subscribers in the world. African markets are expanding nearly twice as fast as Asian markets. The availability of Prepaid or pay as you go services, where the subscriber does not have to commit to a long term contract, has helped fuel this growth on a monumental scale, not only in Africa but on other continents as well.

On a numerical basis, India is the biggest growth market adding about 6 million cell phones every month. With 156.31 million cell phones, teledensity in the country is still low at 17.45% and country expects to reach 500 million subscribers by end of 2010.

All European nations and most Asian and African nations have adopted GSM. In other countries, such as the United States, Australia,India, Japan, and South Korea, legislation does not require any particular standard, and GSM coexists with other standards, such as CDMA and iDEN.

Some cellular systems are pay as you go, where top-ups can be purchased and added to a phone unit, in a wide variety of shops and even ATMs now, so there is no monthly bill. Many are "pay monthly", where a bill is issued every month for the amount of calls and text messages made.

Pay As You Go

Pay As You Go, often shortened to PAYG, is used as a general term for the concept of a prepay mobile phone. In many countries this is referred to as "prepaid" mobile service.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited*****.io***

The concept was initially developed by Eircell in the Republic of Ireland in the 1990s, as a method of letting different types of people (those under the age of 18, those without bank accounts and those without proof of identity) obtain a mobile phone. Originally limited to one TACS handset, costing £99 upfront, the system was an amazing success, despite the high price of calls and a 7p service charge on every operation. The system was branded as Ready To Go, a name still used by Vodafone, who now own Eircell.

A user would buy a phone, usually pre-loaded with some amount of credit, and would purchase extra credit when required. A call cannot be made unless the user has the amount required for that call's minimum charge. Some networks charge more for the credit than you get in call value, often due to service charges and VAT on phone calls; however the opposite is often true, with users paying, for example, €20 for €22 call credit.

The concept has since been copied in many other countries, with virtually every network in every European country supporting it. On many networks, such as Ireland's Meteor, Pay As You Go is the main mode of operation, with account phones being very much second-class. Conversely, in the United States, account phones offer the best features with pay as you go services being far more restricted in functionality.

Often, Pay As You Go customers pay more for their calls and SMS messages, and are limited in what they can do with their phone - calls to international or premium rate numbers may be blocked, and they may not be able to roam. These limitations are often due to the complexity of managing the credit system for high price calls, or when the user was not on their home network.

In the United Kingdom, operators have started placing restrictions on PAYG users (for instance, having a minimum topup amount of more than £10) and raising the price of PAYG phones in an effort to attract users to contract plans (which tend to earn more money than prepay overall).

Culture & Customs

In fewer than twenty years, mobile phones have gone from being rare and expensive pieces of equipment used primarily by the business elite, to a pervasive low-cost personal item. In many countries, mobile phones now outnumber land-line telephones, with most adults and many children now owning mobile phones. In the United States, 50% of children own mobile phones. It is not uncommon for young adults to simply own a mobile phone instead of a land-line for their residence. In some developing countries, where there is little existing fixed-line infrastructure, the mobile phone has become widespread. According to the CIA World Factbook the U.K. now has more mobile phones than people.

With high levels of mobile telephone penetration, a mobile culture has evolved, where the phone becomes a key social tool, and people rely on their mobile phone address book to keep in touch with their friends. Many people keep in touch using SMS, and a whole culture of "texting" has developed from this. The commercial market in SMSs is growing. Many phones even offer Instant Messenger services to increase the simplicity and ease of texting on phones. Mobile phones in Japan, offering Internet capabilities such as NTT DoCoMo's i-mode, offer text messaging via standard e-mail.

The mobile phone itself has also become a totemic and fashion object, with users decorating, customizing, and accessorizing their mobile phones to reflect their personality. This has emerged as its own industry. The sale of commercial ringtones exceeded $2.5 billion in 2004.

Etiquette

Mobile phone etiquette has become an important issue with mobiles ringing at funerals, weddings, cinemas, and plays. Users often speak at increased volume which has led to places like book shops, libraries, movie theatres, doctors' offices, and houses of worship posting signs prohibiting the use of mobile phones, and in some places installing signal-jamming equipment to prevent usage (although in many countries, e.g., the United States, such equipment is currently illegal). Some new buildings such as auditoriums have installed wire mesh in the walls (turning the building into a Faraday cage) which prevents any signal getting through, but does not contravene the jamming laws.

Transportation providers, particularly those involving long-distance services, often offer a "quiet car" where phone use is prohibited, much like the designated non-smoking cars in the past. However many users tend to ignore this as it is rarely enforced, especially if the other cars are crowded and they have no choice but to go in the "quiet car". Mobile phone use on aircraft is also prohibited, because of concerns of possible interference with aircraft radio communications, although the airline Emirates have announced plans to allow limited mobile phone usage on some flights. In a similar vein signs are put up in UK petrol stations prohibiting the use of mobile phones due to hypothetical safety issues. Most schools in the United States have prohibited mobile phones in the classroom due to the high amount of class disruptions that result from their use, and due to the possibility of photographing someone (without consent).

Use In Disaster Response

In Japan, mobile phone companies provide immediate notification of earthquakes and other natural disasters to their customers free of charge. In the event of an emergency, disaster response crews can locate trapped or injured people using the signals from their mobile phones; an interactive menu accessible through the phone's Internet browser notifies the company if the user is safe or in distress.

Use By Drivers

Mobile-phone use while driving is common but controversial. While few jurisdictions have banned motorists from using mobile phones while driving outright, some have banned or restricted drivers from using hand-held mobile phones while exempting phones operated in a hands-free fashion. It is generally agreed that using a hand-held mobile phone while driving is a distraction that brings risk of road traffic accidents. However, some studies have found similarly elevated accident rates among drivers using hands-free phones, suggesting that the distraction of a telephone conversation itself is the main safety problem.

Use of handheld mobile phones by drivers is illegal in many European countries and a number of Asian and South American countries. Use of hands-free mobiles is permitted. However some countries like Japan ban mobile phone use while driving completely. Similar laws exist in six U.S. states with legislation proposed in 40 other states.

Applications

Mobile news services are expanding with many organizations providing "on-demand" news services by SMS. Some also provide "instant" news pushed out by SMS. Mobile telephony also facilitates activism and public journalism being explored by Reuters and Yahoo and small independent news companies such as Jasmine News in Sri Lanka.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

Features

There are many mobile phone features found in today's mobile phones that offer users many more capabilities than only voice calls or text messaging.

General Features

Mobiles are designed to work on cellular networks and contain a standard set of services that allow phones of different types and in different countries to communicate with each other.

Before the phone can be used, a subscription to a mobile phone operator (a.k.a. carrier) is required. For phones on GSM networks, the operator will issue a SIM card which contains the unique subscription and authentication parameters for that customer; alternatively, the carrier will put the customer's handset identifier into its subscriber database so that the handset can make calls on the network. Once the SIM card is inserted into the phone, services can be accessed. Many mobile phones support 'auto-roaming', which permits the same phone to be used in multiple countries. For this to work, the operators of both countries must have a roaming agreement.

Mobile phones include an alarm. Usually it still functions while the phone is turned off or in silent mode.

Mobile phones do not only support voice calls; they can also send and receive data and faxes (if a computer is attached), send short messages (or "text messages"; see SMS), access WAP services, and provide full Internet access using technologies such as GPRS. Mobile phones usually have a clock and a calculator and often one can play some games on them.

Most current models also allow for sending and receiving pictures and have a built-in digital camera. Sound and video recording is often also possible. This feature is generally referred to as MMS. This gives rise to some concern about privacy, in view of possible voyeurism, for example in swimming pools. For this reason, Saudi Arabia has entirely banned the sale of camera phones; South Korea has ordered manufacturers to ensure that all new handsets emit a beep whenever a picture is taken.

GPS receivers are starting to appear integrated or connected (i.e. using bluetooth) to cell phones, primarily to aid in dispatching emergency responders and road tow truck services. This feature is generally referred to as E911.

Push to talk, available on some mobile phones, is a feature that allows the user to talk to another by pressing the button. It functions similar to walkie-talkie.

There are also many features aimed toward personalisation, such as user defined and downloadable ring tones and logos, and interchangeable covers, which have helped in the uptake by the teenage market. Mobile phone content advertising has become massively popular but has also drawn a great deal of criticism. Usually one can choose between a ring tone, a vibrating alert, or a combination of both.

As a result of all these features packed into a tiny device, mobile phones have recently gained reputations for their poor ergonomics. Their small size, plethora of features and modes, and attempts at stylish design often make them difficult and confusing to use.

Multi-Mode Mobile Phones

A multi-mode (also known more specifically as dual, tri or quad band) mobile phone is a phone which is designed to work on more than one GSM radio frequency. The multi-mode case occurs mostly in GSM which was originally specified in the 900 MHz band, but expanded to the 1800 MHz band, later adding 1900 MHz and finally 850 MHz in the Americas. Some multi-mode phones can operate on analog networks as well (for example, dual band, tri-mode: AMPS 800 / CDMA 800 / CDMA 1900).

For a GSM phone, dual-band usually means 850 / 1900 MHz in the United States and Canada, 900 / 1800 MHz in Europe and most other countries. Tri-band means 850 / 1800 / 1900 MHz or 900 / 1800 / 1900 MHz. Quad-band means 850 / 900 / 1800 / 1900 MHz, also called a world phone, since it can work on any GSM network.

For CDMA or TDMA phones, multiple bands usually means it supports both digital and analog communications.

Multi-mode phones have been valuable to enable roaming but are now becoming most important in allowing the introduction of WCDMA without customers having to give up the wide coverage of GSM. Almost every single true 3G phone sold is actually a WCDMA/GSM dual-mode mobile. This is also true of 2.75G phones such as those based on CDMA-2000 or EDGE.

The special challenge involved in producing a multi-mode mobile is in finding ways to share the components between the different standards. Obviously, the phone keypad and display should be shared, otherwise it would be hard to treat as one phone. Beyond that, though, there are challenges at each level of integration. How difficult these challenges are depends on the differences between systems. The different variants of the GSM system have only different frequencies and so aren't even considered true multi-mode phones but rather are called multi-band phones. When talking about IS-95/GSM multi-mode phones, for example, or AMPS/IS-95 phones, the base band processing is very different from system to system. This leads to real difficulties in component integration and so to larger phones.

An interesting special case of multi-mode phones is the WCDMA/GSM phone. The radio interfaces are very different from each other, but mobile to core network messaging has strong similarities, meaning that software sharing is quite easy. Probably more importantly, the WCDMA air interface has been designed with GSM compatibility in mind. It has a special mode of operation, known as punctured mode, in which, instead of transmitting continuously, the mobile is able to stop sending for a short period and try searching for GSM carriers in the area. This mode allows for safe inter-frequency handovers with channel measurements which can only be approximated using "pilot signals" in other CDMA based systems.

A final interesting case is that of mobiles covering the DS-WCDMA and MC-CDMA 3G variants of the CDMA-2000 protocol. Initially, the chip rate of these phones was incompatible. As part of the negotiations related to patents, it was agreed to use compatible chip rates. This should mean that, despite the fact that the air and system interfaces are quite different, even on a philosophical level, much of the hardware for each system inside a phone should be common with differences being mostly confined to software.

As can be deduced from the above, most mobile phone networks now use one of two standards, GSM or CDMA. A third standard, iDEN, is found exclusively in North America and is confined to use by the Nextel network. It's believed that this network will eventually disappear as Nextel merges with Sprint PCS, a CDMA carrier. Similarly, AT&T Wireless's TDMA network in North America is slowly being phased out as a result of its merger with Cingular, a GSM (TDMA) carrier.

Data Communications

Mobile phones are now heavily used for data communications such as SMS messages, browsing mobile web sites, and even streaming audio and video files. The main limiting factors are the size of the screen, lack of a keyboard, processing power and connection speed.

With newer smartphones, screen resolution and processing power has become bigger and better. Some new phone CPUs run at over 400 MHz. Many complex programs are now available for the various smartphones, such as Symbian and Windows Mobile.

Connection speed is based on network support. The data part of the GSM protocol is called GPRS. Many recent GSM phones also support EDGE, a high speed version of GPRS. Some phones also feature real keyboards, such as the Nokia 6820.

As of April 2006, a significant number of models support 3G communications - generally a downlink of up to 384kb/s and an uplink of up to 64kb/s. Recent models include models such as the Nokia 6680 and the Nokia N90. Such phones have access to the Web via a free download of the Opera browser.

New Features

There are also new means of digital communications, such as text messaging and e-mail. As of 2004, even basic phones can send and receive text messages which makes them vulnerable to attack by worms and viruses. Advanced phones capable of e-mail can be susceptible to viruses that can multiply by sending messages through a phone's address book. Of more important concern, a virus may allow unauthorized users to access a phone to find passwords or corporate data stored on the device. Moreover, they can be used to commandeer the phone to make calls or send messages at the owner's expense. Unlike computers that are restricted to only a few widespread operating systems, cellular phones use a variety of systems that require separate programs to be designed in order to disable each one. While reducing overall compatibility from an application design standpoint, this has the beneficial effect of making it harder to design a mass attack. However, the rise of cellular phone operating system programming platforms shared by many manufacturers such as Java, Microsoft operating systems, Linux, or Symbian OS, may in the future change this status quo.

Bluetooth is a wireless communication feature now found in many higher-end phones, and the virus Cabir hijacked this function, sending Bluetooth phones on a search-and-destroy mission to infect other Bluetooth phones. In early November 2004, several web sites began offering a specific piece of software promising ringtones and screensavers for certain phones. Those who downloaded the software found that it turned each icon on the phone's screen into a skull-and-crossbones and disabled their phones, so they could no longer send or receive text messages or access contact lists or calendars. The virus has since been dubbed "Skulls" by security experts. The Commwarrior-A virus was identified in March 2005, and it attempts to replicate itself through MMS to others on the phone's contact list. Like Cabir, Commwarrior-A also tries to communicate via Bluetooth wireless connections with other devices, which can eventually lead to draining the battery. The virus requires user intervention for propagation however. Bluetooth telephones are also subject to bluejacking, which is the generally benign transmission of messages from anonymous Bluetooth users. In 2004, rumors spread of using Bluetooth to arrange casual sex hookups; this activity, widely publicized in both print and online media as toothing, was revealed to be a hoax in 2005.

Video Cameras

As of April 2006, there has been a significant enhancement of the camera capability of mobile phones both for still photograph and video. The Nokia N90 has a 2M pixel camera and can record video at 352x288 pixels and 15 frames per second. The Nokia N93 is reported to provide "DVD quality" video at 30 frames per second. Modern software allows simple editing and publishing of video content from phones, and CNN reports that 80% of consumer video is now shot on phones.

Most people do not walk around with a video camera, but do carry a phone. The arrival of video camera phones is transforming the availability of video to consumers, and is helping to fuel the idea of Citizen journalism.

Forensics & Evidence

Law enforcement globally relies heavily upon mobile telephone evidence. The concerns over terrorism and terrorist use of technology promoted an enquiry by the UK House of Commons Home Affairs Select Committee into the use of evidence from mobile telephone devices, prompting leading mobile telephone forensic specialists to identify forensic techniques available in this area.

Health Impacts

Since the introduction of mobile phones, concerns have been raised about the potential health impacts from cellular phone use. Studies from the National Cancer Institute and researchers at the Danish Institute of Cancer Epidemiology in Copenhagen do not show any link between cellular phone use and cancer. The Danish study only covered analog mobile phone usage up through 1995, and subjects who started mobile phone usage after 1995 were counted as non-users in the study. However, a study by the International Agency for Research on Cancer of 4,500 users found a statistically significant link between tumor frequency and mobile phone use.

Technology

Mobile phones and the network they operate under vary significantly from provider to provider, and nation to nation. However, all of them communicate through electromagnetic microwaves with a cell site base station, the antennas of which are usually mounted on a tower, pole, or building.

The phones have a low-power transceiver that transmits voice and data to the nearest cell sites, usually not more than 5 to 8 miles (approximately 8 to 13 kilometres) away. When the mobile phone or data device is turned on, it registers with the mobile telephone exchange, or switch, with its unique identifiers, and will then be alerted by the mobile switch when there is an incoming telephone call. The handset constantly listens for the strongest signal being received from the surrounding base stations. As the user moves around the network, the mobile device will "handoff" to various cell sites during calls, or while waiting (idle) between calls it will reselect cell sites.

Cell sites have relatively low-power (often only one or two watts) radio transmitters which broadcast their presence and relay communications between the mobile handsets and the switch. The switch in turn connects the call to another subscriber of the same wireless service provider or to the public telephone network, which includes the networks of other wireless carriers. Many of these sites are camouflaged to blend with existing environments, particularly in high-scenery areas.

The dialogue between the handset and the cell site is a stream of digital data that includes digitized audio (except for the first generation analog networks). The technology that achieves this depends on the system which the mobile phone operator has adopted. Some technologies include AMPS for analog, and D-AMPS, CDMA2000, GSM, GPRS, EV-DO, and UMTS for digital communications. Each network operator has a unique radio frequency band.

Mobile Communication Studies

Since 2002, many books have been written on the social impact of mobile phones:

- Agar, Jon, Constant Touch: A Global History of the Mobile Phone, 2004
- Glotz, Peter & Bertsch, Stefan, eds. Thumb Culture: The Meaning of Mobile Phones for Society, 2005
- Katz, James E. & Aakhus, Mark, eds. Perpetual Contact: Mobile Communication, Private Talk, Public Performance, 2002
- Kavoori, Anandam & Arceneaux, Noah, eds. The Cell Phone Reader: Essays in Social Transformation, 2006
- Ling, Rich, The mobile connection 2004
- Ling, Rich and Pedersen, Per, eds. Mobile Communications: Renegotiation of the Social Sphere 2005
- Nyíri, Kristóf, ed. Mobile Communication: Essays on Cognition and Community, 2003
- Nyíri, Kristóf, ed. Mobile Learning: Essays on Philosophy, Psychology and Education, 2003
- Nyíri, Kristóf, ed. Mobile Democracy: Essays on Society, Self and Politics, 2003
- Nyíri, Kristóf, ed. A Sense of Place: The Global and the Local in Mobile Communication, 2005
- Nyíri, Kristóf, ed. Mobile Understanding: The Epistemology of Ubiquitous Communication, 2006
- Levinson, Paul, Cellphone: The Story of the World's Most Mobile Medium, and How It Has Transformed Everything! 2004
- Rheingold, Howard, Smart Mobs: The Next Social Revolution, 2002

Mobile Telephony – IMEI

What Is It?

The International Mobile Equipment Identity is a number unique to every GSM and UMTS mobile phone. It is usually found printed on the phone underneath the battery and can also be found by dialing the sequence *#06# into the phone.

The IMEI number is used by the GSM network to identify valid devices and therefore can be used to stop a stolen phone from accessing the network. For example, if a mobile phone is stolen, the owner can call his or her network provider and instruct them to "ban" the phone using its IMEI number. This renders the phone useless, regardless of whether the phone's SIM is changed.

Unlike the Electronic Serial Number or MEID of CDMA and other wireless networks, the IMEI is only used to identify the device, and has no permanent or semi-permanent relation to the subscriber. Instead, the subscriber is identified by transmission of an IMSI number, which is stored on a SIM card which can (in theory) be transferred to any handset. However, many network and security features are enabled by knowing the current device being used by a subscriber.

The IMEI (14 digits plus check digit) or IMEISV (16 digits) includes information on the origin, model, and serial number of the device. The structure of the IMEI/SV are specified in 3GPP TS 23.003. The model and origin comprise the initial 8-digit portion of the IMEI/SV, known as the Type Allocation Code (TAC). The remainder of the IMEI is manufacturer-defined, with a Luhn check digit at the end (which is never transmitted).

As of 2004, the format of the IMEI is AA-BBBBBB-CCCCCC-D, although it may not always be displayed this way. The IMEISV drops the Luhn check digit in favour of an additional 2 digits for the Software Version Number (SVN) in the format AA-BBBBBB-CCCCCC-EE

- AA - Reporting Body Identifier, indicating the GSMA-approved group that allocated the model TAC
- BBBBBB - The remainder of the TAC
- CCCCCC - Serial sequence of the model
- D - Luhn check digit of the entire number (or zero)
- EE - Software Version Number (SVN).

The Luhn algorithm or Luhn formula, also known as the "modulus 10" or "mod 10" algorithm, is a simple checksum formula used to validate a variety of identification numbers, such as credit card numbers and Canadian Social Insurance Numbers. It was created by IBM scientist Hans Peter Luhn and described in U.S. Patent 2,950,048 , filed on January 6, 1954, and granted on August 23, 1960.

The algorithm is in the public domain and is in wide use today. It is not intended to be a cryptographically secure hash function; it was designed to protect against accidental errors, not malicious attacks. Most credit cards and many government identification numbers use the algorithm as a simple method of distinguishing valid numbers from collections of random digits.

Prior to 2002, the TAC was 6 digits long and followed by a two-digit Final Assembly Code (FAC), which was a manufacturer-specific code indicating the location of the device's construction.

For example the IMEI code 35-209900-176148-1 or IMEISV code 35-209900-176148-23 tells us the following:

- TAC: 352099 so it was issued by the BABT and has the allocation number 2099
- FAC: 00 so it was numbered during the transition phase from the old format to the new format (described below)
- SNR: 176148 - uniquely identifying a unit of this model
- CD: 1 so it is a GSM Phase 2 or higher
- SVN: 23 - The 'software version number' identifying the revision of the software installed on the phone. 99 is reserved.

The format changed from April 1, 2004 when the Final Assembly Code ceased to exist and the Type Approval Code increases to eight digits in length and became known as the Type Allocation Code. From January 1, 2003 until this time the FAC for all phones was 00.

The Reporting Body Identifier is allocated by the Global Decimal Administrator; the first two digits must be decimal (ie less than 0xA0) for it to be an IMEI and not an MEID.

The new CDMA Mobile Equipment Identifier (MEID) uses the same basic format as the IMEI.

Retrieving IMEI Information From A GSM Device

On many devices the IMEI number can be retrieved by entering *#06#. The IMEI number of a GSM device can be retrieved by sending the command AT+CGSN. For more information refer the 3GPP TS 27.007, Section 5.4 /2/ standards document.

Retrieving IMEI Information from an Ericsson or Sony Ericsson handset can be done by entering these keys: Right * Left Left * Left * (Other service menu items will be presented with this key combination).

The IMEI information can be retrieved from most Nokia mobile phones by pressing *#92702689# (*#WAR0ANTY#), this opens the warranty menu in which the first item is the serial number (the IMEI). The warranty menu also shows other information such as the date the phone was made and the life timer of the phone.

The IMEI can frequently be displayed through phone menus, under a section titled 'System Information', 'Device', 'Phone Info' or similar. Many phones also have the IMEI listed on a label in the battery compartment.

IMEI & The Law

Many countries have acknowledged the use of the IMEI in reducing the effect of mobile phone theft, which has increased exponentially over the last few years. For example, in the United Kingdom under the Mobile Telephones (Re-programming) Act, changing the IMEI of a phone, or possessing equipment that can change it, is considered an offence under some circumstances.

There is a misunderstanding amongst some regulators that the existence of a formally allocated IMEI number range to a GSM terminal implies that the terminal is approved or complies with regulatory requirements. This is not the case. The linkage between regulatory approval and IMEI allocation was removed in April 2000 with the introduction of the European R&TTE Directive. Since that date, IMEIs have been allocated by BABT (acting on behalf of the GSM Association) to legitimate GSM terminal manufacturers without the need to provide evidence of approval.

In Singapore, however, despite the high usage of mobile phones and the increasingly frequent cases of mobile phone theft, there is apparently no infrastructure available to "ban" the use of a stolen phone, given the IMEI. None of the local service providers seem to be aware of the IMEI or its use and most service providers believe that most phones still do not possess the technology to allow the service providers to "ban" the use of a stolen phone.

Blacklist Of Stolen Devices

When mobile equipment is stolen or lost, the operator or owner will typically contact the Central Equipment Identity Register (CEIR), which blacklists the device in all operator switches so that it will in effect become unusable, making theft of mobile equipment a useless business.

The IMEI number is not supposed to be easy to change, making the CEIR blacklisting effective. However this is not always the case: IMEI may be easy to change with special tools and operators may even flatly ignore the CEIR blacklist.

Difficulties

- "New IMEIs can be programmed into stolen handsets and 10% of IMEIs are not unique." According to a BT-Cellnet spokesman quoted by the BBC.
- Facilities do not exist to unblock numbers listed in error on all networks. This is possible in the UK, however, where the user who initially blocked the IMEI must quote a password chosen at the time the block was applied.

Mobile Telephony – SIM

What Is It?

A Subscriber Identity Module (SIM) is a removable smart card, available in two standard sizes the first the size of a credit card (85.60 mm × 53.98 mm x 0.76 mm), while its more popular mini version has a width of 25 mm, a height of 15 mm, and a thickness of 0.76 mm. SIM cards store securely the key identifying a mobile phone service subscriber.

The SIM card allows users to change phones easily by removing the SIM card and inserting it into another mobile phone, thereby eliminating the need for activation of the new mobile phone on the network. The use of SIM card is mandatory in the GSM world. The equivalent of a SIM in UMTS is called the Universal Subscriber Identity Module (USIM), whereas the Removable User Identity Module (RUIM) is more popular in the CDMA world.

W-SIM is a SIM card which also integrates core cellular technology into the card itself.

Memory Storage Size

SIMs are available in many storage sizes, the largest being the 512 KB SIM. Smaller sized SIMs such as the 32 KB and 16 KB are most prevalent in area with less developed GSM networks. There are also Large Memory SIMs, in the order of 128 - 512 Mega Bytes.

Operating Systems

SIM operating systems come in two main flavors: Native and Java. Native SIMs are based on proprietary, vendor specific software whereas, the Java SIMs are based on the Java programming language. Java cards have the advantage of being hardware independent and interoperable.

Data

SIM cards store network specific information used to authenticate and identify subscribers on the Network, the most important of these are the ICCID, IMSI, Authentication Key (Ki), Local Area Identity (LAI). The SIM also stores other carrier specific data such as the SMSC (Short Message Service Centre) number, Service Provider Name (SPN), Service Dialling Numbers (SDN), and Value Added Service (VAS) applications.

ICCID

Each SIM is internationally identified by its ICCID (International Circuit Card ID). ICCIDs are stored in the SIM cards and are also engraved or printed on the SIM card body during a process called personalization.

IMSI

SIM cards are identified on their individual operator networks by holding a unique International Mobile Subscriber Identity. Mobile operators connect mobile phone calls and communicate with their market SIM cards using their IMSI.

An International Mobile Subscriber Identity, or IMSI [im-zee], is a unique number that is associated with all GSM and Universal Mobile Telecommunications System (UMTS) network mobile phone users. The number is stored in the Subscriber Identity Module (SIM). It is sent by the mobile phone to the network and is also used to acquire other details of the mobile in the Home Location Register (HLR) or as locally copied in the Visitor Location Register. In order to avoid the subscriber being identified and tracked by eavesdroppers on the radio interface, the IMSI is sent as rarely as possible and a randomly generated TMSI is sent instead.

Also the IMSI is used in ANY mobile network that interconnect with others. We are talking here about CDMA and EVDO networks and this number is provisioned in the phone directly or in the RUIMS card (a CDMA analogue equivalent to a SIM card in GSM)

An IMSI is usually fifteen digits long. However, they can be shorter (eg. MTN South Africa's are 14 digits). The first three digits are the Mobile Country Code, and the next digits are the Mobile Network Code (MNC). The MNC can be either two digits long (normal e.g. in Europe) or three digits long (normal in North America), the remaining digits, up to the maximum length are the mobile subscriber identification number (MSIN) within the network's customer base.

The IMSI conforms to the ITU E.212 numbering standard.

Authentication Key (Ki)

The Ki is a 32 digit number used in authenticating the SIMs on the mobile network. Each SIM holds a unique Ki assigned to it by the operator during the personalization process. The Ki is also stored on a database (known as Home Location Register or HLR) on the carrier’s network.

Authentication Process

1. On mobile startup the SIM sends its IMSI to the Mobile Operator requesting access and authentication.
2. The operator network searches its database for the incoming IMSI and its associated Ki.
3. The operator network then generates a Random Number (Rand) and signs it with the SIM’s Ki computing another number known as Signed Response (SRES_1)
4. The operator network then sends the RAND to the SIM card that also signs it with its Ki and sends the result (SRES_2) back to the operator network.

5. The operator network then compares its computed SRES_1 with the SIMs computed SRES_2. If the two numbers match the SIM is authenticated and granted access to the operator's network.

Location Area Identity

The SIM stores network state information which is broadcast to it from the network, such as the Location Area Identity (LAI). Operators networks are divided into Location Areas, each having a unique LAI number. When the Mobile changes its location from one Location Area to another it stores its new LAI in SIM and sends it to the operator network to inform network with its new location. If the handset is turned off and back on again it will take data off the SIM and search for the LAI it was in. This saves time by avoiding having to search the whole list of frequencies that the telephone normally would.

SIM Locks

A SIM lock, confused with PIN or PUK code, is a capability built-in to GSM phones by mobile phone manufacturers. Network providers use this capability to restrict the use of in these phones to specific countries and network providers. Currently, phones can be locked to only accept SIM cards from one or more of the following:

- Countries (your phone will work in the US, but not the UK)
- Network providers (e.g. T-Mobile, Orange, etc)
- Service providers (another name for network providers)
- SIM types (i.e. only specific SIM cards can be used with the phone).

Virtually all mobile phones are shipped with country and network provider locks. In addition, most modern phones have firmware installed on them which is specific to the network provider. If you have a Cingular phone, it displays the Cingular logo and supports only Cingular features. This firmware is installed by the service provider and is separate from the locking mechanism. You can unlock a Cingular phone to work with a T-Mobile network, but the phone will still display the Cingular logo and only support Cingular features. Phones can be unbranded by uploading a different firmware version, a procedure recommended for advanced users only.

SIM Cloning

SIM cloning consists of illegally duplicating the GSM Subscriber Identity Module (a SIM card must be entered into a GSM phone in order for it to connect to a network, except for the basic GSM emergency 112 calls or local emergency calls like emergency 911 calls if the user is in the United States) identification and placing calls or using other charged services using the account of the cloned SIM.

In the early 1990s, due to poor security, cloning was more common than it is today. Cloning has now been rendered more challenging technically (as physical access to the SIM card is required as opposed to simply being within radio reach).

SIM cloning is now more difficult to perform, as merely duplicating the contents of the SIM does not enable a duplicate SIM to operate, as the SIM itself performs security operations on the data contained inside to thwart such copying. In order to function, the cloned SIM needs to perform security operations on the data contained, just like the 'originial' one.

SIM cloning is a great concern of security/police services as it renders GSM location-based service (LBS), and thus tracking a GSM user, unreliable if more than one handset is using the same SIM.

In the movie The Bourne Supremacy, the main character clones a SIM card to enable him to impersonate a foreign agent. It is widely held that actually cloning a SIM could not be accomplished as quickly as it was in the film.

The security has been improved with the newer generation of SIM cards to be used with 3GSM networks, known as USIMs. The new specification implements a new publicly announced algorithm (the KASUMI algorithm).

Mobile Telephony – GSM

What Is It?

The Global System for Mobile Communications (GSM: originally from Groupe Spécial Mobile) is the most popular standard for mobile phones in the world. GSM service is used by over 2 billion people across more than 212 countries and territories. The ubiquity of the GSM standard makes international roaming very common between mobile phone operators, enabling subscribers to use their phones in many parts of the world. GSM differs significantly from its predecessors in that both signaling and speech channels are Digital call quality, which means that it is considered a second generation (2G) mobile phone system. This fact has also meant that data communication was built into the system from the 3rd Generation Partnership Project (3GPP).

From the point of view of the consumers, the key advantage of GSM systems has been higher digital voice quality and low cost alternatives to making calls such as text messaging. The advantage for network operators has been the ability to deploy equipment from different vendors because the open standard allows easy inter-operability. Like other cellular standards GSM allows network operators to offer roaming services which mean subscribers can use their phones all over the world.

As the GSM standard continued to develop, it retained backward compatibility with the original GSM phones; for example, packet data capabilities were added in the Release '97 version of the standard, by means of GPRS. Higher speed data transmission has also been introduced with EDGE in the Release '99 version of the standard.

History Of GSM

The growth of cellular telephone systems took off in the early 1980s, particularly in Europe. The lack of a technological standardization prompted the European Conference of Postal and Telecommunications Administrations (CEPT) to create the Groupe Spécial Mobile (GSM) in 1982 with the objective of developing a standard for a mobile telephone system that could be used across Europe.

The first GSM network was launched in 1991 by Radiolinja in Finland.

In 1989, GSM responsibility was transferred to the European Telecommunications Standards Institute (ETSI), and phase I of the GSM specifications were published in 1990. By the end of 1993, over a million subscribers were using GSM phone networks being operated by 70 carriers across 48 countries.

Radio Interface

GSM is a cellular network, which means that mobile phones connect to it by searching for cells in the immediate vicinity. GSM networks operate in four different frequency ranges. Most GSM networks operate in the 900 MHz or 1800 MHz bands. Some countries in the Americas (including the United States and Canada) use the 850 MHz and 1900 MHz bands because the 900 and 1800 MHz frequency bands were already allocated.

The rarer 400 and 450 MHz frequency bands are assigned in some countries, notably Scandinavia, where these frequencies were previously used for first-generation systems.

In the 900 MHz band the uplink frequency band is 890-915 MHz, and the downlink frequency band is 935-960 MHz. This 25 MHz bandwidth is subdivided into 124 carrier frequency channels, each spaced 200 kHz apart. Time division multiplexing is used to allow eight full-rate or sixteen half-rate speech channels per radio frequency channel. There are eight radio timeslots (giving eight burst periods) grouped into what is called a TDMA frame. Half rate channels use alternate frames in the same timeslot. The channel data rate is 270.833 kbit/s, and the frame duration is 4.615 ms.

The transmission power in the handset is limited to a maximum of 2 watts in GSM850/900 and 1 watt in GSM1800/1900.

GSM has used a variety of voice codecs to squeeze 3.1kHz audio into between 6 and 13kbps. Originally, two codecs, named after the types of data channel they were allocated, were used, called "Full Rate" (13kbps) and "Half Rate" (6kbps). These used a system based upon linear predictive coding (LPC). In addition to being efficient with bitrates, these codecs also made it easier to identify more important parts of the audio, allowing the air interface layer to prioritize and better protect these parts of the signal.

GSM was further enhanced in 1997 with the GSM-EFR codec, a 12.2kbps codec that uses a full rate channel. Finally, with the development of UMTS, EFR was refactored into a variable-rate codec called AMR-Narrowband, which is high quality and robust against interference when used on full rate channels, and less robust but still relatively high quality when used in good radio conditions on half-rate channels.

There are four different cell sizes in a GSM network - macro, micro, pico and umbrella cells. The coverage area of each cell varies according to the implementation environment. Macro cells can be regarded as cells where the base station antenna is installed on a mast or a building above average roof top level. Micro cells are cells whose antenna height is under average roof top level; they are typically used in urban areas. Picocells are small cells whose diameter is a few dozen meters; they are mainly used indoors. Umbrella cells are used to cover shadowed regions of smaller cells and fill in gaps in coverage between those cells.

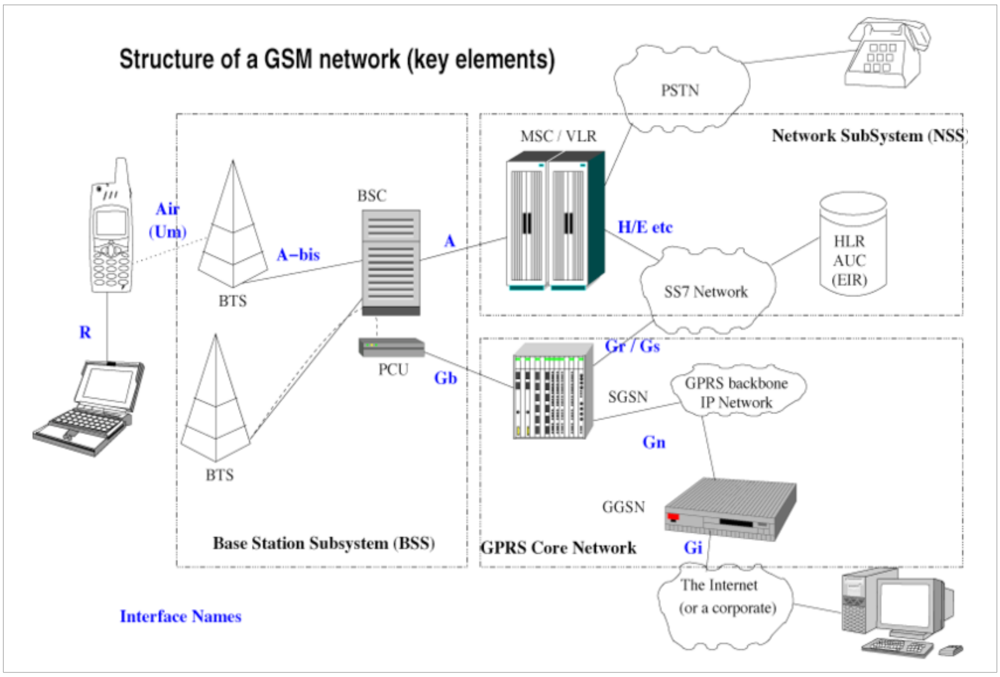
Cell horizontal radius varies depending on antenna height, antenna gain and propagation conditions from a couple of hundred meters to several tens of kilometers. The longest distance the GSM specification supports in practical use is 35 km or 22 miles. There are also several implementations of the concept of an extended cell, where the cell radius could be double or even more, depending on the antenna system, the type of terrain and the timing advance.

Indoor coverage is also supported by GSM and may be achieved by using an indoor picocell base station, or an indoor repeater with distributed indoor antennas fed through power splitters, to deliver the radio signals from an antenna outdoors to the separate indoor distributed antenna system. These are typically deployed when a lot of call capacity is needed indoors, for example in shopping centers or airports. However, this is not a prerequisite, since indoor coverage is also provided by in-building penetration of the radio signals from nearby cells.

The modulation used in GSM is Gaussian minimum shift keying (GMSK), a kind of continuous-phase frequency shift keying. In GMSK, the signal to be modulated onto the carrier is first smoothed with a Gaussian low-pass filter prior to being fed to a frequency modulator, which greatly reduces the interference to neighboring channels (adjacent channel interference).

A nearby GSM handset is usually the source of the "dit dit dit, dit dit dit, dit dit dit" signal that can be heard from time to time on home stereo systems, televisions, computers, and personal music devices. When these audio devices are in the near field of the GSM handset, the radio signal is strong enough that the solid state amplifiers in the audio chain function as a detector. The clicking noise itself represents the power bursts that carry the TDMA signal. These signals have been known to interfere with other electronic devices, such as car stereos and portable audio players. This is a form of RFI, and could be mitigated or eliminated by use of additional shielding and/or bypass capacitors in these audio devices[citation needed], however, the increased cost of doing so is difficult for a designer to justify.

Network Structure



The network behind the GSM system seen by the customer is large and complicated in order to provide all of the services which are required. It is divided into a number of sections and these are each covered in separate articles.

- the Base Station Subsystem (the base stations and their controllers).
- the Network and Switching Subsystem (the part of the network most similar to a fixed network). This is sometimes also just called the core network.
- the GPRS Core Network (the optional part which allows packet based Internet connections).
- all of the elements in the system combine to produce many GSM services such as voice calls and SMS.

Subscriber Identity Module

One of the key features of GSM is the Subscriber Identity Module (SIM), commonly known as a SIM card. The SIM is a detachable smart card containing the user's subscription information and phonebook. This allows the user to retain his or her information after switching handsets. Alternatively, the user can also change operators while retaining the handset simply by changing the SIM. Some operators will block this by allowing the phone to use only a single SIM, or only a SIM issued by them; this practice is known as SIM locking, and is illegal in some countries.

In the United States, Canada, Europe and Australia, many operators lock the mobiles they sell. This is done because the price of the mobile phone is typically subsidised with revenue from subscriptions and operators want to try to avoid subsidising competitor's mobiles. A subscriber can usually contact the provider to remove the lock for a fee, utilize private services to remove the lock, or make use of ample software and websites available on the Internet to unlock the handset themselves. While most web sites offer the unlocking for a fee, some do it for free. The locking applies to the handset, identified by its International Mobile Equipment Identity (IMEI) number, not to the account (which is identified by the SIM card). It is always possible to switch to another (non-locked) handset if such other handset is available.

Some providers will unlock the phone for free if the customer has held an account for a certain period. Third party unlocking services exist that are often quicker and lower cost than that of the operator. In most countries removing the lock is legal. Cingular and T-Mobile provide free unlock services to their customers after 3 months of subscription.[citation needed]

In countries like India, Pakistan, Indonesia, Belgium, etc., all phones are sold unlocked. However, in Belgium, it is unlawful for operators there to offer any form of subsidy on the phone's price. This was also the case in Finland until April 1, 2006, when selling subsidized combinations of handsets and accounts became legal though operators have to unlock phone free of charge after a certain period (at most 24 months).

GSM Security

GSM was designed with a moderate level of security. The system was designed to authenticate the subscriber using shared-secret cryptography. Communications between the subscriber and the base station can be encrypted. The development of UMTS introduces an optional USIM, that uses a longer authentication key to give greater security, as well as mutually authenticating the network and the user - whereas GSM only authenticated the user to the network (and not vice versa). The security model therefore offers confidentiality and authentication, but limited authorization capabilities, and no non-repudiation.

GSM uses several cryptographic algorithms for security. The A5/1 and A5/2 stream ciphers are used for ensuring over-the-air voice privacy. A5/1 was developed first and is a stronger algorithm used within Europe and the United States; A5/2 is weaker and used in other countries. A large security advantage of GSM over earlier systems is that the Ki, the crypto variable stored on the SIM card that is the key to any GSM ciphering algorithm, is never sent over the air interface. Serious weaknesses have been found in both algorithms, and it is possible to break A5/2 in real-time in a ciphertext-only attack. The system supports multiple algorithms so operators may replace that cipher with a stronger one.

Services

GSM Localization

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

GSM localization is the use of multilateration to determine the location of GSM mobile phones, usually with the intent to locate the user.

There are several types of Localization-Based Systems:

- Cell Identification - The accuracy of this method can be as good as a few hundred meters in urban areas, but as poor as 32 km in suburban areas and rural zones. The accuracy depends on the known range of the particular network base station serving the handset at the time of positioning.
- Enhanced Cell Identification - With this method one can get a precision similar to Cell Identification, but for rural areas, with circular sectors of 550 meters.
- TOA - Time of arrival
- AOA - Angle of arrival
- E-OTD - This is similar to TOA, but the position is estimated by the mobile phone, not by the base station. The precision of this method depends on the number of available LMUs in the networks, varying from 50 to 200 m.
- Cell Broadcast - This technology lets mobile operators notify users on their current location. Users' handsets can display the location name (city, country, etc.), the area code or any other location-relevant information.
- Assisted-GPS - A largely GPS-based technology, which uses an operator-maintained ground station to correct for GPS errors caused by the atmosphere/topography. Assisted-GPS positioning technology typically fallsback to cell-based positioning method when indoors or in an urban-canyon environment.

GSM Services

GSM services are a standard collection of applications and features available to mobile phone subscribers all over the world. The GSM standards are defined by the 3GPP collaboration and implemented in hardware and software by equipment manufacturers and mobile phone operators. The common standard makes it possible to use the same phones with different companies' services, or even roam into different countries. GSM is the world's most dominant mobile phone standard.

The design of the service is moderately complex because it must be able to locate a moving phone anywhere in the world, and accommodate the relatively short battery life, limited input/output capabilities, and weak radio transmitters on mobile devices.

GSM supports a comprehensive set of supplementary services that complement and support the telephony and data services described above. They are all defined in GSM standards. A partial listing of supplementary services follows.

- Call Forwarding. This service gives the subscriber the ability to forward incoming calls to another number if the called mobile unit is not reachable, if it is busy, if there is no reply, or if call forwarding is allowed unconditionally.

- Barring of Outgoing Calls. This service makes it possible for a mobile subscriber to prevent all outgoing calls.

- Barring of Incoming Calls. This function allows the subscriber to prevent incoming calls. The following two conditions for incoming call barring exist: baring of all incoming calls and barring of incoming calls when roaming outside the home PLMN.

- Advice of Charge (AoC). The AoC service provides the mobile subscriber with an estimate of the call charges. There are two types of AoC information: one that provides the subscriber with an estimate of the bill and one that can be used for immediate charging purposes. AoC for data calls is provided on the basis of time measurements.

- Call Hold. This service enables the subscriber to interrupt an ongoing call and then subsequently reestablish the call. The call hold service is only applicable to normal telephony.

- Call Waiting. This service enables the mobile subscriber to be notified of an incoming call during a conversation. The subscriber can answer, reject, or ignore the incoming call. Call waiting is applicable to all GSM telecommunications services using a circuit-switched connection.

- Multiparty service. The multiparty service enables a mobile subscriber to establish a multiparty conversation - that is, a simultaneous conversation between three and six subscribers. This service is only applicable to normal telephony.

- Calling Line Identification presentation/restriction. These services supply the called party with the integrated services digital network (ISDN) number of the calling party. The restriction service enables the calling party to restrict the presentation. The restriction overrides the presentation.

- Closed User Groups (CUGs). CUGs are generally comparable to a PBX. They are a group of subscribers who are capable of only calling themselves and certain numbers.

- Explicit Call Transfer (ECT). This service allows a user who has two calls to connect these two calls together and release its connections to both other parties.

MMS

Multimedia Messaging Service (MMS) is a standard for a telephony messaging systems that allow sending messages that includes multimedia objects (images, audio, video, rich text) and not just text messages as in Short message service (SMS). It is mainly deployed in cellular networks along with other messaging systems like SMS, Mobile Instant Messaging and Mobile E-Mail. Its main standardization effort is done by 3GPP, 3GPP2 and Open Mobile Alliance (OMA).

SMS

Short Message Service (SMS) is available on most digital mobile phones and a steadily increasing range of other devices (including Pocket PC, desktop computers and some fixed phones) that permit the sending of short text messages between these devices.

The terms text messaging, text messages, more colloquially SMSes, texts, or even txts and its variants are more commonly used in North America, the UK, Spain and the Philippines, while most other countries prefer the term SMS.

Text messages are also often used to interact with automated systems, such as ordering products and services for mobile phones, or participating in contests. There are some services available on the Internet that allow users to send text messages free of direct charge to the sender, although users in other parts of the world should be aware that users of North American networks will often have to pay to receive any SMS text message.

WAP (Wireless Application Protocol)

Wireless Application Protocol or WAP is an open international standard for applications that use wireless communication. Its principal application is to enable access to the Internet from a mobile phone or PDA.

A WAP browser is to provide all of the basic services of a computer based web browser but simplified to operate within the restrictions of a mobile phone. WAP is now the protocol used for the majority of the world's mobile internet sites, known as WAP sites. The Japanese i-mode system is currently the only other major competing wireless data protocol.

Mobile internet sites, or WAP sites, are websites written in, or dynamically converted to, WML (Wireless Markup Language) and accessed via the WAP browser.

Before the introduction of WAP, service providers had extremely limited opportunities to offer interactive data services. Interactive data applications are required to support now commonplace activities such as:

- email by mobile phone
- tracking of stock market prices
- sports results
- news headlines
- music downloads

WAP: Possible Failure

WAP was hyped at the time of its introduction, leading users to expect WAP to have the performance of the Web. One telco's advertising showed a cartoon WAP user "surfing" through a Neuromancer-like "information space". In terms of speed, ease of use, appearance, and interoperability, the reality fell far short of expectations. This led to the wide usage of sardonic phrases such as "Worthless Application Protocol", "Wait And Pay", and so on.

Critics advanced several explanations for the early failure of WAP. Some are technical criticisms:

- The idiosyncratic WML language, which cut users off from the true HTML Web, leaving only native WAP content and Web-to-WAP "proxified" content available to WAP users. However, others argue that technology at that stage would simply not have been able to give access to anything but custom-designed content.
- Under-specification of terminal requirements. In the early WAP "standards", there were many optional features and under-specified requirements, which meant that compliant devices would not necessarily interoperate properly. This resulted in great variability in the actual behavior of phones. As an example, some phone models would not accept a page more than 1 Kb in size; others would downright crash. The user interface of devices was also underspecified: as an example, accesskeys (e.g., the ability to press '4' to access directly the fourth link in a list) were variously implemented depending on phone models (sometimes with the accesskey number automatically displayed by the browser next to the link, sometimes without it, and sometimes accesskeys were not implemented at all).
- Constrained user interface capabilities. Terminals with small black and white screens and few buttons, as the early WAP terminals were, are not very apt at presenting a lot of information to their user, which compounded the other problems: one would have

had to be extra careful in designing the user interface on such a resource-constrained device.
- Lack of good authoring tools. The problems above might have been alleviated by a WML authoring tool that would have allowed content providers to easily publish content that would interoperate flawlessly with many models, adapting the pages presented to the User-Agent type. However, the development kits which existed did not provide such a general capability. Developing for the web was easy: with a text editor and a web browser, anybody could get started, thanks also to the forgiving nature of most desktop browser rendering engines. By contrast, the stringent requirements of the WML specifications, the variability in terminals, and the demands of testing on various wireless terminals, along with the lack of widely available desktop authoring and emulation tools, considerably lengthened the time required to complete most projects.

Other criticisms are oriented towards the wireless carriers' particular implementations of WAP:

- Neglect of content providers. Some wireless carriers had assumed a "build it and they will come" strategy, meaning that they would just provide the transport of data as well as the terminals, and then wait for content providers to publish their services on the Internet and make their investment in WAP useful. However, content providers received little help or incentive to go through the complicated route of development. Others, notably in Japan (cf. below), had a more thorough dialogue with their content provider community, which was then replicated in modern, more successful WAP services such as i-mode in Japan or the Gallery service in France.
- Lack of openness. Most wireless carriers sold their WAP services that were "open", in that they allowed users to reach any service expressed in WML and published on the Internet. However, they also made sure that the first page that clients accessed was their own "wireless portal", which they controlled very closely. Given the difficulty in typing up fully qualified URLs on a phone keyboard, most users would give up going "off portal"; by not letting third parties put their own entries on the operators' wireless portal, some contend that operators cut themselves off from a valuable opportunity. On the other hand, some operators argue that their customers would have wanted them to manage the experience and, on such a constrained device, avoid giving access to too many services.

WAP: Possible Success

However, WAP has seen huge success in Japan. While the largest operator NTT DoCoMo has famously disdained WAP in favor of its in-house system i-mode, rival operators KDDI (au) and Vodafone Japan have both been successful with the WAP technology. In particular, J-Phone's Sha-Mail picture mail and Java (JSCL) services, as well as (au)'s chakuuta/chakumovie (ringtone song/ringtone movie) services are based on WAP. After being shadowed by the initial success of i-mode, the two smaller Japanese operators have been gaining market share from DoCoMo since spring 2001.

Korea is also leading the world in providing advanced WAP services. WAP on top of the CDMA2000 network has been proven to be the state of the art wireless data infrastructure.

According to the Mobile Data Association, June 2004 saw a considerable increase of 42% in its recorded number of WAP pages viewed compared with the same period in 2003. This took the total for the second quarter of 2004 to 4 billion.

Between 2003 and 2004, WAP made a stronger resurgence with the introduction of Wireless services (such as Vodafone Live!, T-Mobile T-Zones and other easily-accessible services). Operator revenues were generated by transfer of GPRS and UMTS data which is a different model to the Web, and usage was up. People are starting to use WAP and the early failures have been masked, as the real point of the system – access to wireless services and applications – has come to the forefront.

Spin-off technologies, such as MMS (Multimedia Messaging Service) (picture messaging), a combination of WAP and SMS, have further driven the protocol. An enhanced appreciation of device diversity, supported by the concomitant changes to WAP content to be more device-specific rather than being aimed at a lowest common denominator, has allowed for the content presented to be more compelling and usable. As a result, the adoption rate of WAP technology is on the upswing.

GPRS

General Packet Radio Service (GPRS) is a Mobile Data Service available to users of GSM and IS-136 mobile phones. GPRS data transfer is typically charged per megabyte of transferred data, while data communication via traditional circuit switching is billed per minute of connection time, independent of whether the user has actually transferred data or he has been in an idle state. GPRS can be utilized for services such as WAP access, SMS and MMS, but also for Internet communication services such as email and web access. In the future, it is expected that low cost voice over IP will be made available in cell phones.

2G cellular systems combined with GPRS is often described as "2.5G", that is, a technology between the second (2G) and third (3G) generations of mobile telephony. It provides moderate speed data transfer, by using unused TDMA channels in for example the GSM system. Originally there was some thought to extend GPRS to cover other standards, but instead those networks are being converted to use the GSM standard, so that GSM is the only kind of network where GPRS is in use. GPRS is integrated into GSM standards releases starting with Release 97 and onwards. First it was standardized by ETSI but now that effort has been handed onto the 3GPP.

Cell Broadcast

Cell Broadcast (CB) messaging is a mobile technology feature defined by the ETSI's GSM committee and is part of the GSM standard. It is also known as Short Message Service - Cell Broadcast (SMS-CB). Cell Broadcast is designed for simultaneous delivery of messages to multiple users in a specified area. Whereas the Short Message Service - Point to Point (SMS-PP) is a one-to-one and one-to-a-few service, Cell Broadcast is a one-to-many geographically focused messaging service. Cell Broadcast messaging is also supported by UMTS, as defined by 3GPP.

Cell Broadcast messaging was technologically demonstrated in Paris for the first time, in 1997. By now all GSM phones and mobile operator base stations support the feature. Some mobile operators use Cell Broadcast for communicating the area code of the antenna cell to the mobile user (via channel 050), for nationwide or citywide alerting, weather reports, mass messaging, location based news, etc. Not all operators have the Cell Broadcast messaging function activated in their network yet.

Cell Broadcast is a technology that allows a text or binary message to be defined and distributed to all mobile terminals connected to a set of cells. Whereas SMS messages are sent point-to-point, Cell Broadcast messages are sent point-to-area. This means that one Cell

Broadcast message can reach a huge number of terminalsa at once. In other words, Cell Broadcast messages are directed to radio cells, rather than to a specific terminal. A Cell Broadcast message is an unconfirmed push service, meaning that the originator of the message does not know who has received the message, allowing for services based on anonymity. Mobile telephone user manuals describe how the user can switch the receiving of Cell Broadcast messages on or off.

Cell Broadcast messaging has a number of features that make it particularly appropriate for emergency purposes:

- It is not affected by traffic load; therefore, it will be usable during a disaster when load spikes tend to crash networks, as the 7 July 2005 London bombings showed. Another example was during the Tsunami catastrophe in Asia. Dialog GSM, an operator in Sri Lanka was able to provide ongoing emergency information to its subscribers, to warn of incoming waves, to give news updates, to direct people to supply and distribution centres, and even to arrange donation collections using Celltick's Cell Broadcast Center, based on Cell Broadcast Technology. Also it does not itself cause any significant load on the network, so would not add to the problem.

Cell broadcast is not widely deployed today. In the U.S., most handsets do not have cell broadcast capability, and the major U.S. operators have not deployed the technology in their networks.

Cell Broadcast is a mobile technology that allows messages (up to 15 pages of up to 93 characters) to be broadcast to all mobile handsets and similar devices within a designated geographical area. The broadcast range can be varied, from a single cell to the entire network.

This technology is used in deploying location-based subscriber services, such as regional auctions, local weather, traffic conditions and 'nearest' services (like requesting the nearest hospital). It enables messages to be communicated to multiple mobile phone customers who are located within a given part of its network coverage area at the time the message is broadcast.

Comparison Of Mobile Phone Standards

Global System for Mobile Communications (aka GSM) and IS-95 (aka cdmaOne) are the two most prevalent mobile communication technologies. Both technologies have to solve the same problem: to divide the finite RF spectrum among multiple users.

TDMA (Time Division Multiple Access - underlying technology used in GSM's 2G) does it by chopping up the channel into sequential time slices. Each user of the channel takes turns to transmit and receive signals. In reality, only one person is actually using the channel at a specific moment. This is analogous to time-sharing on a large computer server.

CDMA (Code Division Multiple Access - underlying technology used in GSM's 3G and IS-95's 2G) on the other hand, uses a special type of digital modulation called spread spectrum which spreads the voice data over a very wide channel in pseudorandom fashion. The receiver undoes the randomization to collect the bits together and produce the sound.

As a trivial comparison imagine a cocktail party, where couples are talking to each other in a single room. The room represents the available bandwidth. In GSM, a speaker takes turns talking to a listener. The speaker talks for a short time and then stops to let another pair talk. There is never more than one speaker talking in the room, no one has to worry about two conversations mixing. In CDMA, any speaker can talk at any time; however each uses a different language. Each listener can only understand the language of their partner. As more and more couples talk, the background noise (representing the noise floor) gets louder, but because of the difference in languages, conversations do not mix.

Advantages Of 2G GSM

- GSM is mature; this maturity means a more stable network with robust features.
- Less signal deterioration inside buildings.
- Ability to use repeaters.
- Talktime is generally higher in GSM phones due to the pulse nature of transmission.
- The availability of Subscriber Identity Modules allows users to switch networks and handsets at will.
- GSM covers virtually all parts of the world so international roaming is not a problem.
- The much bigger number of subscribers globally creates a better network effect for GSM handset makers, carriers and end users.
- Tools for unlocking GSM phones are widely available. Meaning it is easier to keep the same phone and go with another carrier using the same technology.

Disadvantages Of 2G GSM

- Pulse nature of TDMA transmission used in 2G interferes with some electronics, especially certain audio amplifiers. 3G uses W-CDMA now.
- Intellectual property is concentrated among a few industry participants, creating barriers to entry for new entrants and limiting competition among phone manufacturers.
- GSM has a fixed maximum cell site range of 35 km, which is imposed by technical limitations.

Advantages Of IS-95

- Capacity is IS-95's biggest asset. It can accommodate more users per MHz of bandwidth than any other technology.
- IS-95 has no built-in limit to the number of concurrent users.
- IS-95 uses precise clocks that do not limit the distance a tower can cover.
- IS-95 consumes less power and covers large areas so cell size in IS-95 is larger.
- IS-95 is able to produce a reasonable call with lower signal (cell phone reception) levels.
- IS-95 uses soft handoff, reducing the likelihood of dropped calls.
- IS-95's variable rate voice coders reduce the rate being transmitted when speaker is not talking, which allows the channel to be packed more efficiently.
- Has a well-defined path to higher data rates.

Disadvantages Of IS-95

- Most technologies are patented and must be licensed from Qualcomm.
- Breathing of base stations, where coverage area shrinks under load. As the number of subscribers using a particular site goes up, the range of that site goes down.
- Because IS-95 towers interfere with themselves, they are normally installed on much shorter towers. Because of this, IS-95 may not perform well in hilly terrain.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

- IS-95 covers a smaller portion of the world, and IS-95 phones are generally unable to roam internationally.
- Manufacturers are often hesitant to release IS-95 devices due to the smaller market, so features are sometimes late in coming to IS-95 devices.
- Tools for unlocking CDMA phones are not common. Meaning every time you subscribe with a new provider you must buy a new phone.

3G

3G (or 3-G) is short for third-generation technology. It is used in the context of mobile phone standards. The services associated with 3G provide the ability to transfer simultaneously both voice data (a telephone call) and non-voice data (such as downloading information, exchanging email, and instant messaging). In marketing 3G services, video telephony has often been used as the killer application for 3G.

Worldwide roll-out of 3G networks was delayed in some countries by the enormous costs of additional spectrum licensing fees. In many parts of the world 3G networks do not use the same radio frequencies as 2G, requiring mobile operators to build entirely new networks and license entirely new frequencies; a notable exception is the United States where carriers operate 3G service in the same frequencies as other services. The license fees in some European countries were particularly high, bolstered by initial excitement over 3G's potential. Other delays were as a result of the expenses related to upgrading equipment for the new systems.

Japan and South Korea were relatively quick to adopt 3G, because their governments prioritize technological infrastructure development, and spectrum licensing fees are minimal.

The first country which introduced 3G on a large commercial scale was Japan. In 2005, about 40% of subscribers used 3G networks only, with 2G being on the way out. It was expected that the transition from 2G to 3G would be largely completed during 2006, and upgrades to the next 3.5G stage with 3 Mbit/s data rates were under way.

The successful 3G introduction in Japan showed that video telephony was not the killer application for 3G networks after all. The real-life usage of video telephony on 3G networks was found to be a small fraction of all services. On the other hand, downloading of music found strong acceptance by customers. Music download services in Japan were pioneered by KDDI with the EZchakuuta and Chaku Uta Full services.

3G networks are not IEEE 802.11 networks. IEEE 802.11 networks are short range, higher-bandwidth (primarily) data networks, while 3G networks are wide area cellular telephone networks which evolved to incorporate high-speed internet access and video telephony.

Background

According to the GSA, in December 2005 there were 100 3G networks in operation in 40 countries in the world. In Asia, Europe, and the USA, telecommunication companies use WCDMA technology with the support of around 100 terminal designs to operate 3G mobile networks.

In 2001, NTT DoCoMo—one of the giant telecommunication companies in Japan—was the first telecommunication company to launch a commercial WCDMA network. The introduction of 3G services within Europe began in early 2003.

The official 3G mobile network is the systems and services based on the International Telecommunication Union (ITU) family of standards under the International Mobile Telecommunications programme, "IMT-2000". A boost was given to 3G mobile networks in Europe when the European Union council suggested that the 3G operators should cover 80% of the European national populations by the end of 2005. In Africa, Vodafone Egypt (also known as CLICK GSM) will provide the service in Egypt in the middle of 2007. Early 2007, Vodacom Tanzania switched on its 3G HSPDA in Dar ea salaam. With the installation of a 3G HSDPA network, Tanzania is only the second country in Africa with such technology, the first being South Africa.In March 2007, Nigeria awarded Third Generation (3G) telecommunication licenses to the nation's three major GSM companies and a relatively unknown operator, Alheri Engineering Co. Ltd, to enable them expand their scope of operation in the industry.

Features

The most significant feature offered by third generation (3G) mobile technologies is the capacity to support greater numbers of voice and data customers — especially in urban centres — as well as higher data rates at lower incremental cost than 2G.

By using the radio spectrum in bands identified, which is provided by the ITU for Third Generation IMT-2000 mobile services, it subsequently licensed to operators. 3G uses 5 MHz channel carrier width to deliver significantly higher data rates and increased capacity compared with 2G networks.

The 5 MHz channel carrier provides optimum use of radio resources for operators who have been granted large, contiguous blocks of spectrum. On the other hand, it also helps to reduce the cost to 3G networks while being capable of providing extremely high-speed data transmission to users.

It also allows the transmission of 384kbps for mobile systems and 2Mbps for stationary systems. 3G users are expected to have greater capacity and improved spectrum efficiency, which will allow them to access global roaming between different 3G networks.

Issues

Even though 3G has successfully been introduced to European and Asian mobile users, there are some issues that are debated by 3G providers and users:

- High input fees for the 3G service licenses
- Great differences in the licensing terms
- Current high debt of many telecommunication companies, making it more of a challenge to build the necessary infrastructure for 3G
- Member State support to the financially troubled operators
- Health aspects of the effects of electromagnetic waves
- Expense of 3G phones
- Lack of 2G mobile user buy-in for 3G wireless service
- Lack of coverage because it is still new service
- High prices of 3G mobile services in some countries, including Internet access

E-Commerce – Electronic Commerce

What Is It?

Electronic Commerce is exactly analogous to a marketplace on the Internet. Electronic Commerce (also referred to as EC, e-commerce eCommerce or ecommerce) consists primarily of the distributing, buying, selling, marketing and servicing of products or services over electronic systems such as the Internet and other computer networks. The information technology industry might see it as an electronic business application aimed at commercial transactions; in this context, it can involve electronic funds transfer, supply chain management, e-marketing, online marketing, online transaction processing, electronic data interchange (EDI), automated inventory management systems, and automated data collection systems. Electronic commerce typically uses electronic communications technology of the World Wide Web, at some point in the transaction's lifecycle, although of course electronic commerce frequently depends on computer technologies other than the World Wide Web, such as databases, and e-mail, and on other non-computer technologies, such as transportation for physical goods sold via e-commerce.

E-Commerce according to Person Halls book E-Commerce started in 1994 with the first banner ad being placed on a website.

According to the October 2006 Forrester Research report entitled, "US eCommerce: Five-Year Forecast And Data Overview, "Nontravel online retail revenues will top the quarter-trillion-dollar mark by 2011. The diver of this growth? A segment of the most active Web shopping households that is roughly 8 million strong. This group of consumers is extremely comfortable with technology and values convenience above all else in the online retail experience. As retailers begin to wade through their copious data warehouses and understand the who, what, when, where, why, and how of this segment, they will benefit from targeting these customers."

Historical Development

the meaning of term "electronic commerce" has changed over the last 30 years. Originally, "electronic commerce" meant the facilitation of commercial transactions electronically, usually using technology like Electronic Data Interchange (EDI) and Electronic Funds Transfer (EFT), where both were introduced in the late 1970s, for example, to send commercial documents like purchase orders or invoices electronically.

The 'electronic' or 'e' in e-commerce refers to the technology/systems; the 'commerce' refers to be traditional business models. E-commerce is the complete set of processes that support commercial/business activities on a network. In the 1970s and 1980s, this would also have involved information analysis. The growth and acceptance of credit cards, automated teller machines (ATM) and telephone banking in the 1980s were also forms of e-commerce. However, from the 1990s onwards, this would include enterprise resource planning systems (ERP), data mining and data warehousing.

In the dot com era, it came to include activities more precisely termed "Web commerce" -- the purchase of goods and services over the World Wide Web, usually with secure connections (HTTPS, a special server protocol that encrypts confidential ordering data for customer protection) with e-shopping carts and with electronic payment services, like credit card payment authorizations.

Today, it encompasses a very wide range of business activities and processes, from e-banking to offshore manufacturing to e-logistics. The ever growing dependence of modern industries on electronically enabled business processes gave impetus to the growth and development of supporting systems, including backend systems, applications and middleware. Examples are broadband and fiber-optic networks, supply-chain management software, customer relationship management software, inventory control systems and financial accounting software.

When the Web first became well-known among the general public in 1994, many journalists and pundits forecast that e-commerce would soon become a major economic sector. However, it took about four years for security protocols (like HTTPS) to become sufficiently developed and widely deployed. Subsequently, between 1998 and 2000, a substantial number of businesses in the United States and Western Europe developed rudimentary web sites.

Although a large number of "pure e-commerce" companies disappeared during the dot-com collapse in 2000 and 2001, many "brick-and-mortar" retailers recognized that such companies had identified valuable niche markets and began to add e-commerce capabilities to their Web sites. For example, after the collapse of online grocer Webvan, two traditional supermarket chains, Albertsons and Safeway, both started e-commerce subsidiaries through which consumers could order groceries online.

The emergence of e-commerce also significantly lowered barriers to entry in the selling of many types of goods; accordingly many small home-based proprietors are able to use the internet to sell goods. Often, small sellers use online auction sites such as EBay(tm), or sell via large corporate websites like Amazon.com, in order to take advantage of the exposure and setup convenience of such sites.

Success Factors In E-Commerce

Technical & Organizational Aspects

In many cases, an e-commerce company will survive not only based on its product, but by having a competent management team, good post-sales services, well-organized business structure, network infrastructure and a secured, well-designed website. Such factors include:

1. Sufficient work done in market research and analysis. E-commerce is not exempt from good business planning and the fundamental laws of supply and demand. Business failure is as much a reality in e-commerce as in any other form of business.
2. A good management team armed with information technology strategy. A company's IT strategy should be a part of the business re-design process.
3. Providing an easy and secured way for customers to effect transactions. Credit cards are the most popular means of sending payments on the internet, accounting for 90% of online purchases. In the past, card numbers were transferred securely between the customer and merchant through independent payment gateways. Such independent payment gateways are still used by most small and home businesses. Most merchants today process credit card transactions on site through arrangements made with commercial banks or credit cards companies.
4. Providing reliability and security. Parallel servers, hardware redundancy, fail-safe technology, information encryption, and firewalls can enhance this requirement.

5. Providing a 360-degree view of the customer relationship, defined as ensuring that all employees, suppliers, and partners have a complete view, and the same view, of the customer. However, customers may not appreciate the big brother experience.
6. Constructing a commercially sound business model.
7. Engineering an electronic value chain in which one focuses on a "limited" number of core competencies -- the opposite of a one-stop shop. (Electronic stores can appear either specialist or generalist if properly programmed.)
8. Operating on or near the cutting edge of technology and staying there as technology changes (but remembering that the fundamentals of commerce remain indifferent to technology).
9. Setting up an organization of sufficient alertness and agility to respond quickly to any changes in the economic, social and physical environment.
10. Providing an attractive website. The tasteful use of colour, graphics, animation, photographs, fonts, and white-space percentage may aid success in this respect.
11. Streamlining business processes, possibly through re-engineering and information technologies.
12. Providing complete understanding of the products or services offered, which not only includes complete product information, but also sound advisors and selectors.

Naturally, the e-commerce vendor must also perform such mundane tasks as being truthful about its product and its availability, shipping reliably, and handling complaints promptly and effectively. A unique property of the Internet environment is that individual customers have access to far more information about the seller than they would find in a brick-and-mortar situation. (Of course, customers can, and occasionally do, research a brick-and-mortar store online before visiting it, so this distinction does not hold water in every case.)

Customer-Oriented

A successful e-commerce organization must also provide an enjoyable and rewarding experience to its customers. Many factors go into making this possible. Such factors include:

1. Providing value to customers. Vendors can achieve this by offering a product or product-line that attracts potential customers at a competitive price, as in non-electronic commerce.
2. Providing service and performance. Offering a responsive, user-friendly purchasing experience, just like a flesh-and-blood retailer, may go some way to achieving these goals.
3. Providing an incentive for customers to buy and to return. Sales promotions to this end can involve coupons, special offers, and discounts. Cross-linked websites and advertising affiliate programs can also help.
4. Providing personal attention. Personalized web sites, purchase suggestions, and personalized special offers may go some of the way to substituting for the face-to-face human interaction found at a traditional point of sale.
5. Providing a sense of community. Chat rooms, discussion boards, soliciting customer input and loyalty programs (sometimes called affinity programs) can help in this respect.
6. Owning the customer's total experience. E-tailers foster this by treating any contacts with a customer as part of a total experience, an experience that becomes synonymous with the brand.
7. Letting customers help themselves. Provision of a self-serve site, easy to use without assistance, can help in this respect. This implies that all product information is available, cross-sell information, advise for product alternatives, and supplies & accessory selectors.

8. Helping customers do their job of consuming. E-tailers and online shopping directories can provide such help through ample comparative information and good search facilities. Provision of component information and safety-and-health comments may assist e-tailers to define the customers' job.

Problems

Even if a provider of E-commerce goods and services rigorously follows these "key factors" to devise an exemplary e-commerce strategy, problems can still arise. Sources of such problems include:

1. Failure to understand customers, why they buy and how they buy. Even a product with a sound value proposition can fail if producers and retailers do not understand customer habits, expectations, and motivations. E-commerce could potentially mitigate this potential problem with proactive and focused marketing research, just as traditional retailers may do.
2. Failure to consider the competitive situation. One may have the will to construct a viable book e-tailing business model, but lack the capability to compete with Amazon.com.
3. Inability to predict environmental reaction. What will competitors do? Will they introduce competitive brands or competitive web sites? Will they supplement their service offerings? Will they try to sabotage a competitor's site? Will price wars break out? What will the government do? Research into competitors, industries and markets may mitigate some consequences here, just as in non-electronic commerce.
4. Over-estimation of resource competence. Can staff, hardware, software, and processes handle the proposed strategy? Have e-tailers failed to develop employee and management skills? These issues may call for thorough resource planning and employee training.
5. Failure to coordinate. If existing reporting and control relationships do not suffice, one can move towards a flat, accountable, and flexible organizational structure, which may or may not aid coordination.
6. Failure to obtain senior management commitment. This often results in a failure to gain sufficient corporate resources to accomplish a task. It may help to get top management involved right from the start.
7. Failure to obtain employee commitment. If planners do not explain their strategy well to employees, or fail to give employees the whole picture, then training and setting up incentives for workers to embrace the strategy may assist.
8. Under-estimation of time requirements. Setting up an e-commerce venture can take considerable time and money, and failure to understand the timing and sequencing of tasks can lead to significant cost overruns. Basic project planning, critical path, critical chain, or PERT analysis may mitigate such failings. Profitability may have to wait for the achievement of market share.
9. Failure to follow a plan. Poor follow-through after the initial planning, and insufficient tracking of progress against a plan can result in problems. One may mitigate such problems with standard tools: benchmarking, milestones, variance tracking, and penalties and rewards for variances.
10. Becoming the victim of organized crime. Many syndicates have caught on to the potential of the Internet as a new revenue stream. Two main methods are as follows: (1) Using identity theft techniques like phishing to order expensive goods and bill them to some innocent person, then liquidating the goods for quick cash; (2) Extortion by using a network of compromised "zombie" computers to engage in distributed denial of service attacks against the target Web site until it starts paying protection money.

11. Failure to expect the unexpected. Too often new businesses do not take into account the amount of time, money or resources needed to complete a project and often find themselves without the necessary components to become successful.

Product Suitability

Certain products or services appear more suitable for online sales; others remain more suitable for offline sales.

Many successful purely virtual companies deal with digital products, (including information storage, retrieval, and modification), music, movies, office supplies, education, communication, software, photography, and financial transactions. Examples of this type of company include: Google, eBay and Paypal. Other successful marketers such as use Drop shipping or Affiliate marketing techniques to facilitate transactions of tangible goods without maintaining real inventory. Examples include Amazing Refund and numerous sellers on eBay.

Virtual marketers can sell some non-digital products and services successfully. Such products generally have a high value-to-weight ratio, they may involve embarrassing purchases, they may typically go to people in remote locations, and they may have shut-ins as their typical purchasers. Items which can fit through a standard letterbox — such as music CDs, DVDs and books — are particularly suitable for a virtual marketer, and indeed Amazon.com, one of the few enduring dot-com companies, has historically concentrated on this field.

Products such as spare parts, both for consumer items like washing machines and for industrial equipment like centrifugal pumps, also seem good candidates for selling online. Retailers often need to order spare parts specially, since they typically do not stock them at consumer outlets -- in such cases, e-commerce solutions in spares do not compete with retail stores, only with other ordering systems. A factor for success in this niche can consist of providing customers with exact, reliable information about which part number their particular version of a product needs, for example by providing parts lists keyed by serial number.

Purchases of pornography and of other sex-related products and services fulfill the requirements of both virtuality (or if non-virtual, generally high-value) and potential embarrassment; unsurprisingly, provision of such services has become the most profitable segment of e-commerce.

There are also many disadvantages of e-commerce, one of the main ones is fraud. This is where your details (name, bank card number, age, national insurance number) are entered into what look to be a safe site but really it is not. These details can then be used to steal money from you and can be used to buy things on line that you are completely unaware of until it is too late. If this information is leaked into the wrong hands. People are able to steal your identity, and commit more fraud crimes under your name. Finally there are many problems with e commerce some of which are:

Failure to understand customers, why they buy and how they buy. Even a product with a sound value proposition can fail if producers and retailers do not understand customer habits, expectations, and motivations. E-commerce could potentially mitigate this potential problem with proactive and focused marketing research, just as traditional retailers may do. Failure to consider the competitive situation. One may have the will to construct a viable book e-tailing business model, but lack the capability to compete with Amazon. Inability to predict environmental reaction. What will competitors do? Will they introduce competitive brands or competitive web sites? Will they supplement their service offerings? Will they try to sabotage a competitor's site? Will price wars break out? What will the government do? Research into competitors, industries and markets may mitigate some consequences here, just as in non-electronic commerce. Over-estimation of resource competence. Can staff, hardware, software, and processes handle the proposed strategy? Have e-tailer's failed to develop employee and management skills? These issues may call for thorough resource planning and employee training.

Products less suitable for e-commerce include products that have a low value-to-weight ratio, products that have a smell, taste, or touch component, products that need trial fittings — most notably clothing — and products where colour integrity appears important. Nonetheless, Tesco.com has had success delivering groceries in the UK, albeit that many of its goods are of a generic quality, and clothing sold through the Internet is big business in the U.S. Also, the recycling program Cheapcycle sells goods over the Internet, but avoids the low value-to-weight ratio problem by creating different groups for various regions, so that shipping costs remain low.

Acceptance

Consumers have accepted the e-commerce business model less readily than its proponents originally expected. Even in product categories suitable for e-commerce, electronic shopping has developed only slowly. Several reasons might account for the slow uptake, including:

- Concerns about security. Many people will not use credit cards over the Internet due to concerns about theft and credit card fraud.
- Lack of instant gratification with most e-purchases (non-digital purchases). Much of a consumer's reward for purchasing a product lies in the instant gratification of using and displaying that product. This reward does not exist when one's purchase does not arrive for days or weeks.
- The problem of access to web commerce, mainly for poor households and for developing countries. Low penetration rates of Internet access in some sectors greatly reduces the potential for e-commerce.
- The social aspect of shopping. Some people enjoy talking to sales staff, to other shoppers, or to their cohorts: this social reward side of retail therapy does not exist to the same extent in online shopping.
- Poorly designed, bug-infested e-Commerce web sites that frustrate online shoppers and drive them away.
- Inconsistent return policies among e-tailers or difficulties in exchange/return.

Suppliers Offering Financial Services To Electronic Commerce Practitioners

- iBill
- Moneybookers
- PayPal
- WebMoney
- Yahoo!
- Google Checkout
- Paymate
- PaidByCash
- eWise
- VeriSign

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

Biggest Five Online E-Commerce Sites 2006

- Ebay.com
- Yahoo.com
- Amazon.com
- Google.com
- Buy.com

E-Commerce – Online Banking

What Is It?

Online banking (or Internet banking) is a term used for performing transactions, payments etc. over the Internet through a bank, credit union or building society's secure website. This allows customers to do their banking outside of bank hours and from anywhere where Internet access is available. In most cases a web browser is utilized and any normal internet connection is suitable. No special software or hardware is usually needed.

Features

Online banking usually offers such features as:

- Bank statements, with the possibility to import data in a personal finance program such as Quicken or Microsoft Money
- Electronic bill payment
- Funds transfer between a customer's own checking and savings accounts, or to another customer's account
- Investment purchase or sale
- Loan applications and transactions, such as repayments
- Account aggregation to allow the customers to monitor all of their accounts in one place whether they are with their main bank or with other institutions.

There are a growing number of so-called virtual banks that operate exclusively online. These online banks have low costs compared to traditional banks and so they often offer higher interest rates.

Security

Protection through single password authentication, as is the case in most secure Internet shopping sites, is not considered secure enough for personal online banking applications in some countries. Online banking user interfaces are secure sites (generally employing the https protocol) and traffic of all information - including the password - is encrypted, making it next to impossible for a third party to obtain or modify information after it is sent. However, encryption alone does not rule out the possibility of hackers gaining access to vulnerable home PCs and intercepting the password as it is typed in (keylogging). There is also the danger of password cracking and physical theft of passwords written down by careless users.

Many online banking services therefore impose a second layer of security. Strategies vary, but a common method is the use of transaction numbers, or TANs, which are essentially single use passwords. Another strategy is the use of two passwords, only random parts of which are entered at the start of every online banking session. This is however slightly less secure than the TAN alternative and more inconvenient for the user. A third option, used in many European countries and currently being trialled in the UK is providing customers with security token devices capable of generating single use passwords unique to the customer's token (this is called two-factor authentication or 2FA). Another option is using digital certificates, which digitally sign or authenticate the transactions, by linking them to the physical device (e.g. computer, mobile phone, etc). While most online banking in the United States still uses single password protection, the FDIC has issued regulations requiring that banks implement more secure authentication mechanisms by the end of 2006.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

Banks in many European countries (including the Scandinavian countries, the Netherlands, Austria and Belgium) are offering online banking for e-commerce payments directly from customer to merchants.

Fraud

Some customers avoid online banking as they perceive it as being too vulnerable to fraud. The security measures employed by most banks can never be completely safe, but in practice the number of fraud victims due to online banking is very small. This is probably due to the fact that a relatively small number of people use internet banking compared with the total number of banking customers world wide. Indeed, conventional banking practices may be more prone to abuse by fraudsters than online banking. Credit card fraud, signature forgery and identity theft are far more widespread "offline" crimes than malicious hacking. Bank transactions are generally traceable and criminal penalties for bank fraud are high. Online banking becomes less secure if users are careless, gullible or computer illiterate. An increasingly popular criminal practice to gain access to a user's finances is phishing, whereby the user is in some way persuaded to hand over their password(s) to a fraudster.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

E-Commerce – Electronic Funds Transfer (EFT)

What Is It?

Electronic funds transfer or EFT refers to the computer-based systems used to perform financial transactions electronically.

The term is used for a number of different concepts:

- cardholder-initiated transactions, where a cardholder makes use of a payment card
- electronic payments by businesses, including salary payments
- electronic check (or cheque) clearing

Card-Based EFT

EFT may be initiated by a cardholder when a payment card such as a credit card or debit card is used. This may take place at an automated teller machine (ATM) or point of sale (EFTPOS), or when the card is not present, which covers cards used for mail order, telephone order and internet purchases.

Card-based EFT transactions are often covered by the ISO 8583 standard.

Transaction Types

A number of transaction types may be performed, including the following:

- Sale: where the cardholder pays for goods or service.
- Refund: where a merchant refunds an earlier payment made by a cardholder.
- Withdrawal: the cardholder withdraws funds from their account, e.g. from an ATM. The term Cash Advance may also be used, typically when the funds are advanced by a merchant rather than at an ATM.
- Deposit: where a cardholder deposits funds to their own account (typically at an ATM).
- Cashback: where a cardholder withdraws funds from their own account at the same time as making a purchase.
- Inter-account transfer: transferring funds between linked accounts belonging to the same cardholder)
- Payment: transferring funds to a third party account
- Inquiry: a transaction without financial impact, for instance balance inquiry, available funds inquiry, linked accounts inquiry, or request for a statement of recent transactions on the account.
- Administrative: this covers a variety of non-financial transactions including PIN change.

The transaction types offered depend on the terminal. An ATM would offer different transactions from a POS terminal, for instance.

Authorization

EFT transactions require communication between a number of parties. When a card is used at a merchant or ATM, the transaction is first routed to an acquirer, then through a number of networks to the issuer where the cardholder's account is held.

A transaction may be authorized offline by any of these entities through a stand-in agreement. Stand-in authorization may be used when a communication link is not available, or simply to save communication cost or time. Stand-in is subject to the transaction amount being below agreed limits. These limits are calculated based on the risk of authorizing a transaction offline, and thus vary between merchants and card types. Offline transactions may be subject to other security checks such as checking the card number against a 'hotcard' (stolen card) list, velocity checks (limiting the number of offline transactions allowed by a cardholder) and random online authorization.

A transaction may be authorized via a pre-authorization step, where the merchant requests the issuer to reserve an amount on the cardholder's account for a specific time, followed by completion, where the merchant requests an amount blocked earlier with a pre-authorization. This transaction flow in two steps is often used in businesses such as hotels and car rental where the final amount is not known, and the pre-authorization is made based on an estimated amount. Completion may form part of a settlement process, typically performed at the end of the day when the day's completed transactions are submitted.

Authentication

EFT transactions may be accompanied by methods to authenticate the card and the cardholder. The merchant may manually verify the cardholder's signature, or the cardholder's Personal identification number (PIN) may be sent online in an encrypted form for validation by the card issuer. Other information may be included in the transaction, some of which is not visible to the cardholder (for instance magnetic stripe data), and some of which may be requested from the cardholder (for instance the cardholder's address or the CVV2 value printed on the card).

EMV cards are smartcard-based payment cards, where the smartcard technology allows for a number of enhanced authentication measures.

E-Commerce – Bank Account

What Is It?

A bank account is a financial account with a banking institution, recording the financial transactions between the customer and the bank and the resulting financial position of the customer with the bank.

Bank accounts may have a positive, or debit balance, where the bank owes money to the customer; or a negative, or credit balance, where the customer owes the bank money.

Broadly, accounts opened with the purpose of holding credit balances are referred to as deposit accounts; whilst accounts opened with the purpose of holding debit balances are referred to as loan accounts.

Some accounts are defined by their function rather than nature of the balance they hold. Bank accounts designed to process large numbers of transactions may offer credit and debit facilities and therefore do not sit easily with a polarised definition. These transactional accounts are called by different names in different countries: in the U.S. and Canada, they are called "checking accounts"; in the UK, they are termed "current accounts".

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

E-Commerce – Savings Account

What Is It?

Savings accounts are accounts maintained by retail financial institutions that pay interest but can not be used directly as money (for example, by writing a cheque). These accounts let customers set aside a portion of their liquid assets while earning a monetary return.

Overview

Savings accounts are offered by commercial banks, savings and loan associations, credit unions, building societies and mutual savings banks.

Obtaining funds held in a savings account may not be as convenient as from a demand account. For example, one may need to visit an ATM or bank branch, instead of writing a check or using a debit card. However, this transference is easy enough that savings accounts are often termed "near money".

Some savings accounts require funds to be kept on deposit for a minimum length of time, but most permit unlimited access to funds. In the US, Regulation D, 12 CFR 204.2(d)(2) limits the withdrawals, payments, and transfers that a savings account may perform. Banks comply with these regulations differently; some will immediately prevent the transfer from happening, while others will allow the transfer to occur but will notify the account holder upon violation of the regulation. True savings accounts do not offer cheque-writing privileges, although many institutions will call their higher-interest demand accounts or money market accounts "savings accounts."

All savings accounts offer itemized lists of all financial transactions, traditionally through a passbook, but also through a bank statement.

About 65% of people in the United States have savings accounts.

Regulations

In the United States, under Regulation D, 12 CFR 204.2(d)(2), the term "savings deposit" includes a deposit or an account that meets the requirements of Sec. 204.2(d)(1) and from which, under the terms of the deposit contract or by practice of the depository institution, the depositor is permitted or authorized to make up to six transfers or withdrawals per month or statement cycle of at least four weeks. The depository institution may authorize up to three of these six transfers to be made by check, draft, debit card, or similar order drawn by the depositor and payable to third parties. There is no regulation limiting number of deposits, however some banks may choose to limit deposits themselves.

Within most European countries, interest paid on deposit accounts is taxed at source. The high rates of some countries has led to the development of a significant offshore savings industry. The European Union Savings Directive has made arrangements with many offshore financial centres for either information on interest earned to be shared with EU tax authorities or for withholding tax to be deducted on interest paid on offshore accounts, because of concerns relating to potential tax evasion. Account holders must either pay the withholding tax or disclose account holder information to relevant tax authorities.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

Costs

Withdrawals from a savings account are occasionally costly and are sometimes much higher and more time-consuming than the same financial transaction being performed on a demand account. However, most savings accounts do not limit withdrawals, unlike certificates of deposit. In the United States, violations of Regulation D often involve a service charge, or even a downgrade of the account to a checking account. With online accounts, the main penalty is the time required for the Automated Clearing House to transfer funds from the online account to a "brick and mortar" bank where it can be easily accessed. During the period between when funds are withdrawn from the online bank and transferred to the local bank, no interest is earned.

E-Commerce – Transactional Account

What Is It?

A transactional account (North America: checking account or chequing account, United Kingdom and some other countries: current account or cheque account) is a deposit account held at a bank or other financial institution, for the purpose of securely and quickly providing frequent access to funds on demand, through a variety of different channels. Because money is available on demand these accounts are also referred to as demand accounts or demand deposit accounts.

Transactional accounts are meant neither for the purpose of earning interest nor for the purpose of savings, but for convenience of the business or personal client; hence they tend not to bear interest. Instead, a customer can deposit or withdraw any amount of money any number of times, subject to availability of funds.

Features And Access

All transactional accounts offer itemized lists of all financial transactions, either through a bank statement or a passbook. A transactional account allows the account holder to make or receive payments by:

- cash money (coins and banknotes)
- cheque and money order (paper instruction to pay)
- giro (funds transfer, direct deposit)
- direct debit (pre-authorized debit)
- standing order (automatic funds transfer)
- ATM card or debit card (cashless direct payment at a store or merchant)
- SWIFT: International account to account transfer.

Overdrafts

An overdraft occurs when withdrawals from a bank account exceed the available balance. This gives the account a negative balance and in effect means the account provider is providing credit. If there is a prior agreement with the account provider for an overdraft facility, and the amount overdrawn is within this authorised overdraft, then interest is normally charged at the agreed rate. If the balance exceeds the agreed facility then fees may be charged and a higher interest rate might apply.

Cost

The policy of charging a fee for doing financial transactions depends on a variety of factors, including the country and its overall interest rates for lending and for saving, as well as the size of the financial institution and the number of channels of access it offers. This is why virtual banks, operating few or no branches can afford to offer low-cost or free banking, and why, in some countries, transaction fees do not exist, but extremely high lending rates are the norm.

Financial transaction fees may be charged either per item or for a flat rate covering a certain number of transactions (usually charged on a monthly basis). Often, youths, students, senior citizens or high-valued customers do not pay fees for basic financial transactions. Some will offer free transactions for maintaining a very high average balance in their account. Other service charges are applicable for overdraft, non-sufficient funds, the use of an external interbank network, etc. In countries where there are no service charges for transaction fees, there are, on the other hand, other recurring service charges such as a debit card annual fee.

Interest

Unlike savings accounts, for which the primary reason for depositing money is to generate interest, the main function of a transactional account is transactional. Therefore, most providers either pay no interest or pay a low level of interest on credit balances.

E-Commerce – Financial Transaction

What Is It?

A financial transaction is an event or condition under the contract between a buyer and a seller to exchange an asset for payment. In accounting, it is recognized by an entry in the books of account. It involves a change in the status of the finances of two or more businesses or individuals.

Purchase

The most common type of financial transaction. An item or good is exchanged for money. This transaction results in a decrease in the finances of the purchaser and an increase in the benefits of the sellers. An example is a Real estate transaction.

The Loan

A slightly more complicated transaction in which the lender gives a single large amount of money to the borrower now in return for many smaller repayments of the borrower to the lender over time, usually on a fixed schedule. The smaller delayed repayments usually add up to more than the first large amount. The difference in payments is called interest.Here money is not given for any specific reasons.

Mortgage

A combination loan and purchase. A lender gives a large amount of money to a borrower for the specific purpose of purchasing a very expensive item (most often a house). As part of the transaction, the borrower usually agrees to give the item (or some other high value item) to the lender if the loan is not paid back on time. This guarantee of repayment is known as collateral.

Bank Account

A bank is a business that is based almost entirely on financial transactions. In addition to acting as a lender for loans and mortgages, banks act as a borrower in a special type of loan called an account. The lender is known as a customer and gives unspecified amounts of money to the bank for unspecified amounts of time. The bank agrees to repay any amount in the account at any time and will pay small amounts of interest on the amount of money that the customer leaves in the account for a certain period of time. In addition, the bank guarantees that the money will not be stolen while it is in the account, and will reimburse the customer if it is. In return, the bank gets to use the money for other financial transactions as long as they hold it.

Credit-Card Purchase

A special combination of purchase and loan. The seller gives the buyer the good or item as normal, but the buyer pays the seller using a credit card. In this way, the buyer is paying with a loan from the credit card company, usually a bank. The bank or other financial institution issues credit cards to buyers that allow any number of loans up to a certain cumulative amount. Repayment terms for credit card loans, or debts vary, but the interest is often extremely high. An example of common repayment terms would be a minimum payment of the greater of $10 or 3% every month, and a 15-20% interest charge for any unpaid loan amount. In addition to interest, buyers are sometimes charged a yearly fee to use the credit card.

In order to collect the money for their item, the seller must apply to the credit card company with a signed receipt. Sellers usually apply for many payments at regular intervals. The seller is also charged a fee by the credit card company for the privilege of accepting that brand of credit card for purchases. The fee is normally 1-3% of the purchase price.

Thus, in a credit card purchase, the transfer of the item is immediate, but all payments are delayed.

Debit-Card Purchase

This is a special type of purchase. The item or good is transferred as normal, but the purchaser uses a debit card instead of money to pay. A debit card contains an electronic record of the purchaser's account with a bank. Using this card, the seller is able to send an electronic signal to the buyer's bank for the amount of the purchase,and that amount of money is simultaneously debited from the customer's account and credited to the account of the seller. This is possible even if the buyer or seller use different financial institutions. Currently, fees to both the buyer and seller for the use of debit cards are fairly low because the banks want to encourage the use of debit cards. The seller must have a card reader set up in order for such purchases to be made. Debit cards allow a buyer to have access to all the funds in his account without having to carry the money around. It is more difficult to steal such funds than cash, but is still done.

E-Commerce – Issuing Bank

What Is It?

An issuing bank is a bank that offers card association branded payment cards directly to consumers.

The issuing bank assumes primary liability for the consumer's capacity to pay off debts they incur with their card.

In the case of credit cards, the issuing bank extends a line of credit to the consumer. Liability for non-payment is then shared by the issuing bank and the acquiring bank, according to rules established by the card association brand.

Worldwide, over 1.5 billion payment cards are in circulation

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

E-Commerce – Acquiring Bank

What Is It?

An acquiring bank (or acquirer) is the bank or financial institution that accepts payments for the products or services on behalf of a merchant. The term acquirer indicates that the bank accepts or acquires transactions performed using a credit card issued by a bank other than itself.

Merchant Accounts

The acquiring bank contract with the merchant is informally referred to as a merchant account. The arrangement is in fact a line of credit and not a bank account. Under the contract, the acquiring bank exchanges funds with issuing banks on behalf of the merchant, and pays the merchant for the net balance of their daily payment card activity: gross sales, minus reversals, interchange fees, and acquirer fees.

Interchange fees are fixed rates set by the card association, varying by merchant industry. Acquirer fees are an additional markup added to association fees by the acquiring bank, varying at the acquirer's discretion.

Acquirer Risk

The acquiring bank accepts the risk that the merchant will remain solvent over time, and thus has an incentive to take a keen interest in the merchant's products and business practices. Crucial to maintaining an ongoing positive balance is the limiting of reversals of funds. Consumers may trigger the reversal of funds in two ways:

- A card refund is the return of funds to the consumer, voluntarily initiated by the merchant.
- A card chargeback is the return of funds to the consumer, forcibly initiated by the consumer's issuing bank.

Card associations consider a participating merchant to be a risk if more than 1% of payments received result in a chargeback. Visa and MasterCard levy fines against acquiring banks that retain merchants with high chargeback frequency. To defray the cost of any fines received, the acquiring banks are inclined (but not required) to pass such fines on to the merchant.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

E-Commerce – Billing Mediation Platform

What Is It?

A billing mediation platform is a system used to convert datatypes of a certain type to other datatypes, usually for billing purposes. They are used mostly by telephone companies, who typically need to process UDRs (Usage Detail Records). In call scenarios it is known as CDR (Call Detail Record), and among broadband carriers it's referred to as IPDR.

This CDR/UDR data formats will consist of data like NPX,NPA, Call Duaration,peak time flag,call length and all this data may be represented in binary formats.Mediation Manager will read this data and converts into common normalized format.

Billing Systems and all other Downstreams systems will convert again this data to its own understandable format.

Billing mediation platforms get their name from their behavior; they "mediate" between a variety of other systems. In the typical telephone company scenario, the upstream systems (those providing data to the mediation platform) are network elements, such as telephone switches, and the downstream systems (those receiving data from the mediation platform) perform accounting, auditing, archiving, or bill-generation functions. The mediation system collects, collates and prepares data for consumption by the downstream systems, which often accept data only in a limited set of formats.

Typically a mediation platform is used for the following tasks:

1. Collection and validation of CDRs
2. Filtering out of non billing-relevant CDRs
3. Collating
4. Correlation of different input sources CDRs
5. Aggregation of partial CDRs related to the same call
6. Format change and CDRs normalization
7. Business transformation of data

In a telecom billing scenario, mediation is the first step after receiving a CDR. The mediated CDR is forwarded to a rating engine, which calculates the charge associated with the CDRs. In today's world Rating Engines are more becoming necessary for the telecom billing system to meet the growing variant customer needs for different services.

Despite the name, not all of the data transferred via billing mediation platforms is actually used for billing purposes. For instance, the mediation software might generate traffic volume statistics based on the number and origin of the records passing through it. Those statistics could then be used for capacity planning, as part of a network monitoring procedure, or for any other business intelligence applications.

Sophisticated Billing Mediation software from various providers serves end to end functionality for Telecom Operators. Mediation Software Performs various operation from Collection to Downstream Distribution to Modules like Retail Billing, Interconnect Settlement, Business intelligence, Fraud Detection, Revenue Assurance, Test Call Generation. Following list of activities provides an insight on Mediation software activities

1. Collection and Archive

2. Decoding/Encoding
3. CDR Normalization (Common Format)
4. Filtering
5. Conversion
6. Validation
7. Record Enrichment (Using Complex Reference Data)
8. Duplicate Record Detection
9. Aggregation or Correlation
10. Buffering
11. Cloning
12. Sorting
13. Downstream Format Mapping
14. Header and Trailer generation
15. Downstream Distribution
16. Error Messaging and Alarms
17. Auditing and Reports
18. Reconciliation
19. Reference Data Configuration
20. Provisioning services for the subscription.

Complementary to Billing Mediation functions, comprehensive mediation platforms also provide functionality dedicated to Service Provisioning (the two areas frequently intermix as services configured and used by the end customer result in usage data records generation in the network).

Mediating between the systems is not the only job that Mediation Platform can do. Actually this can be used as a provisioning agent. The basic provisioning commands can be configured within the mediation system and whenever we get a request for the system which does the provisioning, the request can be converted into a file, in which mediation can append the service provisioning commands and send it to HLR for activating any request. This of course, load dependent but can come very handy when there is a crisis in the other system.

At core Mediation involves data transfer between various systems with or without modification of data starting Network elements to OSS/BSS systems.Mediation platform for Telecom Practice supports various systems

- Retail Billing
- Wholesale Billing - National and International
- Network Traffic Management
- Dataware Housing
- Reconciliation system
- Fraud Management
- Provisioning Feed to Sub-systems

Telecom operators offer Voice, video, data, fax and internet services to subscribers and partners on various product lines.Mediation products are tuned to provide solutions for complex business challenges.

Some of the main suppliers of Billing Mediation solutions are the following:

- Ericsson (Multi Mediation)
- InTec (Inter-mediatE)

- Comverse (ComverseONE)

E-Commerce – Interbank Network

What Is It?

An interbank network, also known as an ATM consortium or ATM network, is a computer network that connects the ATMs of different banks and permits these ATMs to interact with the ATM cards of non-native banks.

While interbank networks provide capabilities for all ATM cards within the same network to use other banks' ATMs that belong to the same network, the services vary. For instance, when a person uses their ATM card at an ATM that does not belong to their bank, the basic services, such as balance inquiries and withdrawals, are usually available. However, special services, such as the purchase of mobile phone airtime, are usually not accessible to ATM cardholders of banks other than the ATM cardholders of the acquirer (the bank that owns the ATM). Furthermore, many banks will charge a fee to users of cards that do not come from their own bank (in addition to any fees imposed by the bank of the card the person is using).

Interbank networks are convenient because people can access the ATMs of other banks who are members of the network when their own bank's ATM is unavailable. Such is especially convenient for travelers traveling abroad, where multinational interbank networks, like PLUS or Cirrus, are usually available.

Interbank networks also, through different means, permit the use of ATM cards at a point of sale through the use of a special EFTPOS terminal where ATM cards are treated as debit cards.

E-Commerce – Cash Management

What Is It?

In United States banking, cash management, or treasury management, is a marketing term for certain services offered primarily to larger business customers. It may be used to describe all bank accounts (such as checking accounts) provided to businesses of a certain size, but it is more often used to describe specific services such as cash concentration, zero balance accounting, and automated clearing house facilities. Sometimes, private banking customers are given cash management services.

Cash Management Services Generally Offered

The following is a list of services generally offered by banks and utilised by larger businesses and corporations:

- **Account Reconcilement Services:** Balancing a checkbook can be a difficult process for a very large business, since it issues so many checks it can take a lot of human monitoring to understand which checks have not cleared and therefore what the company's true balance is. To address this, banks have developed a system which allows companies to upload a list of all the checks that they issue on a daily basis, so that at the end of the month the bank statement will show not only which checks have cleared, but also which have not. More recently, banks have used this system to prevent checks from being fraudulently cashed if they are not on the list, a process known as positive pay.

- **Advanced Web Services:** Most banks have an Internet-based system which is more advanced than the one available to consumers. This enables managers to create and authorize special internal logon credentials, allowing employees to send wires and access other cash management features normally not found on the consumer web site.

- **Armored Car Services:** Large retailers who collect a great deal of cash may have the bank pick this cash up via an armored car company, instead of asking its employees to deposit the cash.

- **Automated Clearing House:** Services are usually offered by the cash management division of a bank. The Automated Clearing House is an electronic system used to transfer funds between banks. Companies use this to pay others, especially employees (this is how direct deposit works). Certain companies also use it to collect funds from customers (this is generally how automatic payment plans work). This system is criticized by some consumer advocacy groups, because under this system banks assume that the company initiating the debit is correct until proven otherwise.

- **Balance Reporting Services:** Corporate clients who actively manage their cash balances usually subscribe to secure web-based reporting of their account and transaction information at their lead bank. These sophisticated compilations of banking activity may include balances in foreign currencies, as well as those at other banks. They include information on cash positions as well as 'float' (e.g., checks in the process of collection). Finally, they offer transaction-specific details on all forms of

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

payment activity, including deposits, checks, wire transfers in and out, ACH (automated clearinghouse debits and credits), investments, etc.

- **Cash Concentration Services:** Large or national chain retailers often are in areas where their primary bank does not have branches. Therefore, they open bank accounts at various local banks in the area. To prevent funds in these accounts from being idle and not earning sufficient interest, many of these companies have an agreement set with their primary bank, whereby their primary bank uses the Automated Clearing House to electronically "pull" the money from these banks into a single interest-bearing bank account.

- **Lockbox Services:** Often companies (such as utilities) which receive a large number of payments via checks in the mail have the bank set up a post office box for them, open their mail, and deposit any checks found. This is referred to as a "lockbox" service.

- **Positive Pay:** Positive pay is a service whereby the company electronically shares its check register of all written checks with the bank. The bank therefore will only pay checks listed in that register, with exactly the same specifications as listed in the register (amount, payee, serial number, etc.). This system dramatically reduces check fraud.

- **Sweep Accounts:** are typically offered by the cash management division of a bank. Under this system, excess funds from a company's bank accounts are automatically moved into a money market mutual fund overnight, and then moved back the next morning. This allows them to earn interest overnight. This is the primary use of money market mutual funds.

- **Zero Balance Accounting:** can be thought of as somewhat of a hack. Companies with large numbers of stores or locations can very often be confused if all those stores are depositing into a single bank account. Traditionally, it would be impossible to know which deposits were from which stores without seeking to view images of those deposits. To help correct this problem, banks developed a system where each store is given their own bank account, but all the money deposited into the individual store accounts are automatically moved or swept into the company's main bank account. This allows the company to look at individual statements for each store. U.S. banks are almost all converting their systems so that companies can tell which store made a particular deposit, even if these deposits are all deposited into a single account. Therefore, zero balance accounting is being used less frequently.

- **Wire Transfer:** A wire transfer is an electronic transfer of funds. Wire transfers can be done by a simple bank account transfer, or by a transfer of cash at a cash office. Bank wire transfers are often the most expedient method for transferring funds between bank accounts. A bank wire transfer is a message to the receiving bank requesting them to effect payment in accordance with the instructions given. The message also includes settlement instructions. The actual wire transfer itself is virtually instantaneous, requiring no longer for transmission than a telephone call.

- **Controlled Disbursement:** This is another product offered by banks under Cash Management Services. The bank provides a daily report, typically early in the day, that provides the amount of disbursements that will be charged to the customer's account. This early knowledge of daily funds requirement allows the customer to invest any surplus in intraday investment opportunities, typically money market investments. This is different from delayed disbursements, where payments are issued through a remote branch of a bank and customer is able to delay the payment due to increased float time.

In the past, other services have been offered the usefulness of which has diminished with the rise of the Internet. For example, companies could have daily faxes of their most recent transactions or be sent CD-ROMs of images of their cashed checks.

Cash management services can be costly but usually the cost to a company is outweighed by the benefits: cost savings, accuracy, efficiencies, etc.

E-Commerce – Point Of Sale

What Is It?

Point of sales (POS) or checkout refers to both a checkout counter in a shop, and the location where a transaction occurs. Colloquially, a "checkout" refers to a POS terminal or more generally to the hardware and software used for checkouts, the equivalent of an electronic cash register. A POS terminal manages the selling process by a salesperson accessible interface. The same system allows the creation and printing of the voucher.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

E-Commerce – Automated Teller Machine

What Is It?

An automated teller machine (ATM) is a computerized telecommunications device that provides the customers of a financial institution with access to financial transactions in a public space without the need for a human clerk or bank teller. On most modern ATMs, the customer is identified by inserting a plastic ATM card with a magnetic stripe or a plastic smartcard with a chip that contains a unique card number and some security information, such as an expiration date or CVVC (CVV). Security is provided by the customer entering a personal identification number (PIN).

Using an ATM, customers can access their bank accounts in order to make cash withdrawals (or credit card cash advances) and check their account balances as well as purchasing mobile cell phone prepaid credit. ATMs are known by various other names including automated transaction machine,[1] automated banking machine, money machine, bank machine, cash machine, hole-in-the-wall, cashpoint, Bancomat (in various countries in Europe and Russia), Multibanco (after a registered trade mark, in Portugal), and Any Time Money (in India).

Financial Networks

Most ATMs are connected to interbank networks, enabling people to withdraw and deposit money from machines not belonging to the bank where they have their account or in the country where their accounts are held (enabling cash withdrawals in local currency). Some examples of interbank networks include PULSE, PLUS, Cirrus, Interac, Interswitch, STAR, and LINK.

ATMs rely on authorization of a financial transaction by the card issuer or other authorizing institution via the communications network. This is often performed through an ISO 8583 messaging system.

Many banks charge ATM usage fees. In some cases, these fees are charged solely to users who are not customers of the bank where the ATM is installed; in other cases, they apply to all users. Where machines make a charge some people will not use them, but go to a system without fees.

In order to allow a more diverse range of devices to attach to their networks, some interbank networks have passed rules expanding the definition of an ATM to be a terminal that either has the vault within its footprint or utilizes the vault or cash drawer within the merchant establishment, which allows for the use of a scrip cash dispenser.

ATMs typically connect directly to their ATM Controller via either a dial-up modem over a telephone line or directly via a leased line. Leased lines are preferable to POTS lines because they require less time to establish a connection. Leased lines may be comparatively expensive to operate versus a POTS line, meaning less-trafficked machines will usually rely on a dial-up modem. That dilemma may be solved as high-speed Internet VPN connections become more ubiquitous. Common lower-level layer communication protocols used by ATMs to communicate back to the bank include SNA over SDLC, TC500 over Async, X.25, and TCP/IP over Ethernet.

In addition to methods employed for transaction security and secrecy, all communications traffic between the ATM and the Transaction Processor may also be encrypted via methods such as SSL.

Alternative Uses

Although ATMs were originally developed as just cash dispensers, they have evolved to include many other bank-related functions. In some countries, especially those which benefit from a fully integrated cross-bank ATM network (e.g.: Multibanco in Portugal), ATMs include many functions which are not directly related to the management of one's own bank account, such as:

- Deposit currency recognition, acceptance and recycling
- Paying routine bills, fees, and taxes (utilities, phone bills, social security, legal fees, taxes, etc.)
- Printing bank statements
- Updating passbooks
- Loading monetary value into stored value cards
- Purchasing (Postage stamps, Lottery tickets, Train tickets, Concert tickets, Movie tickets, Shopping mall gift certificates.
- Games and promotional features
- Donating to charities
- Cheque Processing Module
- Adding pre-paid cell phone credit.

Increasingly banks are seeking to use the ATM as a sales device to deliver pre approved loans and targeted advertising using products such as ITM (the Intelligent Teller Machine) from CR2 or Aptra Relate from NCR. ATMs can also act as an advertising channel for companies to advertise their own products or third-party products and services.

Manufacturers have demonstrated and have deployed several different technologies on ATMs that have not yet reached worldwide acceptance, such as:

- Biometrics, where authorization of transactions is based on the scanning of a customer's fingerprint, iris, face, etc. Biometrics on ATMs can be found in Asia.
- Cheque/Cash Acceptance, where the ATM accepts and recognises cheques and/or currency without using envelopes. Expected to grow in importance in the US through Check 21 legislation.
- Bar code scanning
- On-demand printing of "items of value" (such as movie tickets, traveler's cheques, etc.)
- Dispensing additional media (such as phone cards)
- Co-ordination of ATMs with mobile phones
- Customer-specific advertising
- Integration with non-banking equipment

Related Devices

A Talking ATM is a type of ATM that provides audible instructions so that persons who cannot read an ATM screen can independently use the machine. All audible information is delivered privately through a standard headphone jack on the face of the machine. Alternatively, some banks such as the Nordea and Swedbank use a built-in external speaker which may be invoked by pressing the talk button on the keypad. Information is delivered to the customer either through pre-recorded sound files or via text-to-speech speech synthesis.

A postal interactive kiosk may also share many of the same components as an ATM (including a vault), but only dispenses items relating to postage.

A scrip cash dispenser may share many of the same components as an ATM, but lacks the ability to dispense physical cash and consequently requires no vault. Instead, the customer requests a withdrawal transaction from the machine, which prints a receipt. The customer then takes this receipt to a nearby sales clerk, who then exchanges it for cash from the till.

A Teller Assist Unit may also share many of the same components as an ATM (including a vault), but they are distinct in that they are designed to be operated solely by trained personnel and not the general public, they do not integrate directly into interbank networks, and are usually controlled by a computer that is not directly integrated into the overall construction of the unit.

E-Commerce – ISO 8583

What Is It?

ISO 8583 Standard for Financial Transaction Card Originated Messages - Interchange message specifications is the International Organization for Standardization standard for systems that exchange electronic transactions made by cardholders using payment cards.

A card-based transaction typically travels from a transaction acquiring device, such as a point-of-sale terminal or an ATM, through a series of networks, to a card issuing system for authorization against the card holder's account. The transaction data contains information derived from the card (e.g., the account number), the terminal (e.g., the merchant number), the transaction (e.g., the amount), together with other data which may be generated dynamically or added by intervening systems. The card issuing system will either authorize or decline the transaction and generate a response message which must be delivered back to the terminal in a timely manner.

ISO 8583 defines a message format and a communication flow so that different systems can exchange these transactions. The vast majority of transactions made at Automated Teller Machines use ISO 8583 at some point in the communication chain, as do transactions made when a customer uses a card to make a payment in a store. In particular, both the MasterCard and Visa networks base their authorization communications on the ISO 8583 standard, as do many other institutions and networks.

Cardholder-originated transactions include purchase, withdrawal, deposit, refund, reversal, balance inquiry, payments and inter-account transfers. ISO 8583 also defines system-to-system messages for secure key exchanges, reconciliation of totals, and other administrative purposes.

Although ISO 8583 defines a common standard, it is not typically used directly by systems or networks. Instead, each network adapts the standard for its own use with custom fields and custom usages.

The placement of fields in different versions of the standard varies; for example, the currency elements of the 1987 and 1993 versions are no longer used in the 2003 version, which holds currency as a sub-element of any financial amount element. As of writing, ISO 8583:2003 has yet to achieve wide acceptance.

An ISO 8583 message is made of the following parts:

- Message Type Indicator (MTI)
- One or more bitmaps, indicating which data elements are present
- Data elements, the fields of the message

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited*****.io***

E-Commerce – Card Security Code (CVV)

What Is It?

The Card Security Code (CSC), sometimes called Card Verification Value or Code (CVV or CVC), is a security feature for credit or debit card transactions, giving increased protection against credit card fraud.

There are actually two security codes:-

- The first code, called CVC1 or CVV1, is encoded on the magnetic stripe of the card and used for transactions in person.

- The second code, and the most cited, is CVV2 or CVC2. This CSC (also known as a CCID or Credit Card ID) is often asked for by merchants so that they can secure "card not present" transactions occurring over the Internet, by mail, fax or over the phone. In many countries in Western Europe, due to increased attempts at card fraud, it is now mandatory to provide this code when the cardholder is not present in person.

This latter CSC should not be confused with the standard card account number appearing in embossed digits. (The standard card number undergoes a separate validation algorithm called the Luhn algorithm, which serves to determine whether or not a given card's number is appropriate.)

Location Of CVV2

The CVV2 is a 3- or 4-digit value printed on the card or signature strip, but not encoded on the magnetic stripe.

- MasterCard, Visa and Discover credit and debit cards have a 3-digit code, called the "CVC2" (card validation code), "CVV2" (card verification value), and "CID" (card identification number), respectively. It is not embossed like the card number, and is always the final group of numbers printed on the back signature panel of the card. New North American MasterCard and Visa cards feature the "CVC2" in a separate panel to the right of the signature strip. This has been done to prevent overwriting of the numbers by signing the card.

- American Express cards have a 4-digit code printed on the front side of the card above the number, referred to as the CID. It is printed flat, not embossed like the card number.

The number is generated when the card is issued, by encrypting the card number and expiry date under a key known only to the issuing bank. Supplying this code in a transaction is intended to verify that the customer has the card in their physical possession.

Security Benefits Of CVV2

Since the CVV2 is not contained on the magnetic stripe of the card, it is not typically included in the transaction when the card is used face to face at a merchant. However, while some merchants in North America, such as Sears and Staples, have only recently begun requiring the code. For American Express cards, this has been invariable practice (for "card not present" transactions) in European Union (EU) states like Ireland and the United Kingdom since the start of 2005. This provides a level of protection to the cardholder, in that a corrupt merchant cannot simply capture the magnetic stripe details of a card and use them later for "card not present" purchases over the phone, mail order or Internet. To do this, a merchant would also have to note the CVV2 visually and record it, which is more likely to arouse the cardholder's suspicion.

Online merchants who require the CVV2 in their transactions are forbidden in the USA by Visa from storing the CVV2 once the individual transaction is authorised and completed. This way, if a database of transactions is compromised, the CVV2 is not included, and the stolen card numbers are less useful.

CVV2 Limitations

The use of the CVV2 cannot protect against phishing scams, where the cardholder is tricked into entering the CVV2 among other card details via a fraudulent website. The growth in phishing has reduced the real-world effectiveness of the CVV2 as an anti-fraud device. There is now also a scam where a phisher has already obtained the card account number (perhaps by hacking a merchant database or from a poorly designed receipt) and gives this information to the victims (lulling them into a false sense of security) before asking for the CVV2 (which is all that the phisher needs).

Since the CVV2 may not be stored by the merchant for any length of time (after the original transaction in which the CVV2 was quoted and then authorized and completed), a merchant who needs to regularly bill a card for a regular subscription would not be able to provide the code after the initial transaction.
Some card issuers do not yet use CVV2 - although MasterCard started in 1997 and Visa in the USA had them issued by 2001.

This means the use of CVV2 codes must remain optional; however, transactions without CVV2 are likely to be subjected to more stringent fraud screening, and fraudulent transactions without CVV2 are more likely to be resolved in favour of the cardholder.

E-Commerce – Electronic Money

What Is It?

Electronic money (also known as electronic cash, electronic currency, digital money, digital cash, digital currency or scrip) refers to money which is exchanged only electronically. Typically, this involves use of computer networks, the internet and digital stored value systems. Electronic Funds Transfer (EFT) and direct deposit are examples of electronic money. Also, it is a collective term for financial cryptography and technologies enabling it.

While electronic money has been an interesting problem for cryptography (see for example the work of David Chaum and Markus Jakobsson), to date, use of digital cash has been relatively low-scale. One rare success has been Hong Kong's Octopus card system, which started as a transit payment system and has grown into a widely used electronic cash system. Another success is Canada's Interac network, which in 2000 at retail (in Canada) surpassed cash [1] as a payment method.

Singapore has a very successful electronic money implementation for its public transportation system (commuter trains, bus, etc), which is very similar to Hong Kong's Octopus card and based on the same type of card (FeliCa). The electronic money, known as EZ-Link by most Singaporeans, is a card the size of an ordinary credit card; it has a smart chip plus a wireless communication module. Passengers just need to tap the EZ-Link when they board the bus and tap the card again when they alight; the bus fare system automatically deducts the calculated bus fare from the EZ-Link value. Recently, McDonalds is setting up EZ-Link payment infrastructure at their fast-food branches all over Singapore's main island. It is believed that in the near future EZ-Link will gain more acceptance as a convenient electronic money solution in Singapore.

Alternative Systems

Technically electronic or digital money is a representation, or a system of debits and credits, used (but not limited to this) to exchange value, within another system, or itself as a stand alone system, online or offline. Also sometimes the term electronic money is used to refer to the provider itself. A private currency may use gold to provide extra security, such as digital gold currency. An electronic currency can be fully backed by gold, like e-gold, or non-gold backed, like eeeCurrency. There are also such e-currencies that are both gold- and non-gold backed, like Webmoney.

Many systems will sell their electronic currency directly to the end user, such as Paypal or e-Bullion, but other systems, such as e-gold or WebMoney, sell only through third party digital currency exchangers, like OmniPay, IceGold or Goldtotem who service orders manually, or automated websites.

In the case of Octopus Card in Hong Kong, deposits work similarly to banks'. After Octopus Card Limited receives money for deposit from users, the money is deposited into banks, which is similar to debit-card-issuing banks redepositing money at central banks.

Some community currencies, like some LETS systems, work with electronic transactions. Cyclos Software allows creation of electronic community currencies.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

The Ripple monetary system (http://ripple.sourceforge.net/) is a project to develop a distributed system of electronic money independent of local currency.

Virtual Debit Cards

Various companies now sell VISA, Mastercard or Maestro debit cards, which can be recharged via electronic money systems. This system has the advantage of greater privacy if a card provider is located offshore, and greater security since the client can never be debited more than the value on the prepaid card. Also debit cards are useful for people who do not have a bank account or are living in countries which do not authorize international money transfers. Generally cards can be recharged with either e-gold, e-Bullion, WebMoney,via a wire transfer, or for U.S. consumers, cash (currency) with product like PaidByCash.

Advantages

Most money in today's world is electronic, and tangible cash is becoming less frequent. With the introduction of internet / online banking, debit cards, online bill payments and internet business, paper money is becoming a thing of the past.

Banks now offer many services whereby a customer can transfer funds, purchase stocks, contribute to their retirement plans (such as Canadian RRSP) and offer a variety of other services without having to handle physical cash or checks. Customers do not have to wait in lines; this provides a lower-hassle environment.

Debit cards and online bill payments allow immediate transfer of funds from an individual's personal account to a business's account without any actual paper transfer of money. This offers a great convenience to many people and businesses alike.

Disadvantages

Although there are many benefits to digital cash, there are also many significant disadvantages. These include fraud, failure of technology, possible tracking of individuals and loss of human interaction.

Fraud over digital cash has been a pressing issue in recent years. Hacking into bank accounts and illegal retrieval of banking records has led to a widespread invasion of privacy and has promoted identity theft.

There is also a pressing issue regarding the technology involved in digital cash. Power failures, loss of records and undependable software often cause a major setback in promoting the technology.

Privacy questions have also been raised; there is a fear that the use of debit cards and the like will lead to the creation by the banking industry of a global tracking system. Some people are working on anonymous ecash to try to address this issue.

Future Evolution

The main focuses of digital cash development are 1) being able to use it through a wider range of hardware such as secured credit cards; and 2) linked bank accounts that would generally be used over an internet means, for exchange with a secure micropayment system such as in large corporations (PayPal).

Furthering network evolution in terms of the use of digital cash, a company named DigiCash is at the focus of creating an e-cash system that would allow issuers to sell electronic coins at some value. When they are purchased they come under someone's own name and are stored on his computer or under his online identity. At all times, the e-cash is linked to the e-cash company and all transactions go through it, so the e-cash company secures anything that is purchased. Only the company knows your information and will properly direct purchases to your location.

Theoretical developments in the area of decentralized money are underway that may rival traditional, centralized money. Systems of accounting such as Altruistic Economics are emerging that are entirely electronic, and can be more efficient and more realistic because they do not assume a zero-sum transaction model.

E-Commerce – Mobile Commerce

What Is It?

Mobile Commerce (also known as M-Commerce, mCommerce or U-Commerce, owing to the ubiquitous nature of its services) is the ability to conduct commerce, using a mobile device e.g. a mobile phone (or cell phone), a PDA, a smartphone while on the move.

Mobile commerce is currently mainly used for the sale of mobile phone ring-tones and games, although as 3G/UMTS services roll out it is increasingly used to enable payment for location-based services such as maps, as well as video and audio content, including full length music tracks. Other services include the sending of information such as football scores via SMS.

Currently the main payment methods used to enable mobile commerce are:

- premium-rate calling numbers,
- charging to the mobile telephone user's bill or
- deducting from their calling credit, either directly or via reverse-charged SMS.

Mobile commerce was coined in the late 1990s during the dot-com boom. The idea that highly profitable mobile commerce applications would be possible though the broadband mobile telephony provided by 2.5G and 3G cellphone services was one of the main reasons for hundreds of billions of dollars in licensing fees paid by European telecommunications companies for UMTS and other 3G licenses in 2000 and 2001.

Other examples of mobile commerce applications are information-on-demand systems like news services or stock tickers, banking and stock brokerage applications by SMS, WAP or iMode.

Trends In Mobile Commerce

PDA’s and cellular phones have become so popular that many businesses are beginning to use m-commerce as a more efficient method of reaching the demands of their customers. Although most trends and advances are seen in Asia and in Europe, North America (Canada and the United States) is also beginning to take advantage of m-commerce.

Banks and other financial institutions are exploring the use of M-Commerce to broaden/retain their business by allowing their customers to not only access account information, e.g. bank balances, stock quotes and financial advice, from anywhere, but also the possibility to make transactions, e.g. purchasing stocks, remitting money, via mobile phones. This service is often referred to as Mobile Banking or M-Banking. The stock market services offered via mobile devices have also become more popular and are known as Mobile Brokerage, as they allow the subscriber to react to market developments in a timely fashion and irrespective of their physical location.

News information is also becoming more popular with subscriptions to daily headlines from anywhere in the world being transmitted to mobile devices. Sports and entertainment are areas that have also grown with the demand for mobile related services. Shopping and reservation services are now more accessible when using mobile devices. Corporations are now using m-commerce to expand everything from services to marketing and advertisement. Although there are currently very few regulations on the use and abuses of mobile commerce, this will change in the next few years. With the increased use of m-commerce comes increased security. Cell phone companies are now spending more money to protect their customers and their information from online intrusions and hackers.

Future Implications

Financial Institutions such as Banks see mobile commerce as offering new channels of service to customers as well as offering them new and innovative products. These financial institutions are working to design and implement new applications that will offer mobile payment (ie. being able to pay for groceries) and mobile brokering. The travel industry, in realizing the possible benefits of m-commerce, is working on technologies that will take care of travel arrangements, update customers on flight status, notify them when this information changes and will offer to make new arrangements based on preset user preferences requiring no input from the user. Therefore, a customer’s entire trip can be scheduled and maintained using only their mobile device. The retail sector is also looking into the possibility of using mobile commerce for making the purchase of merchandize easier. Customers will be able to browse and order products while using a cheaper more secure payment method. An example of this is; instead of using paper catalogues, retailers can send customers a list of products that the customer would be interested in, directly to their mobile device. Additionally, retailers will also be able to track customers at all times and notify them of discounts at local stores in which that customer would be interested in. Shopping will also be easier. Soon, phones will be equipped with “bar-code scanners” and shoppers could scan an item and find out its pricing and availability. In the entertainment industry, m-commerce could be used for the purchasing of movie tickets, verify someone's ID or authorize their reservation information. This industry will also be able to promote wireless gaming and music.

E-Commerce – Mobile Banking

What Is It?

Mobile Banking refers to "provision and availment of banking and financial services with the help of mobile telecommunication devices. The scope of offered services may include facilities to conduct bank and stock market transactions, to administer accounts and to access customized information." [Buse/Tiwari, 2006]

Mobile Banking (also called as M-Banking or mBanking by some) consists of three inter-related applications:

- Mobile Accounting
- Mobile Brokerage
- Mobile Financial Information Services

Most services in Accounting and Brokerage are transaction based. The non-transaction based services of informational nature are however imperative to conduct transaction. For instance, balance enquiries might be needed before committing a money remittance. The accounting and brokerage services are therefore offered invariably in combination with information services. Information services, on the other hand, may be offered as an independent module.

Trends In Mobile Banking

The advancement of Internet has revolutionized the way the financial services industry conducts business. It has empowered organizations with new business models and new ways to offer 24x7 accessibility to their customers. The ability to offer financial transactions online has created new players in the financial services industry, such as online banks, online brokers and wealth managers who offer personalized services.

Over the last few years, the mobile and wireless market has been one of the fastest growing and most interesting markets in the world. It is still growing at a rapid pace. A recent study done by In-Stat/MDR claims that the number of mobile subscribers worldwide will reach 2 billion before the end of 2007. Mobile users have just started to fully utilize the data capabilities in their mobile phones. In Asian countries like India, China, Indonesia and Philippines, where mobile infrastructure is comparatively better than the fixed-line infrastructure and in European countries where mobile phone penetration is very high (80% of consumers use a mobile phone), mobile phone banking is likely to appeal even more.

This opens up huge market for financial institutions interested in offering value added services. With mobile technology, banks can offer a wide range of services to their customers such as doing funds transfer while traveling, receiving online updates of stock price or even performing stock trading while being stuck in traffic. According to the German mobile operator Mobilcom, mobile banking will be the killer application for the next generation of mobile technology.

Mobile devices, especially smartphones, are the most promising way to reach the masses and to create “stickiness” among current customers, due to their ability to provide services anytime, anywhere, high rate of penetration and potential to grow. According to Gartner, shipment of smartphones is growing fast, and should top 20 million units in 2006 alone.

In the last 4 years, banks across the globe have invested billions of dollars to build sophisticated internet banking capabilities. As the trend is shifting to mobile banking, there is a challenge for CIOs and CTOs of these banks to decide on how to leverage their investment in internet banking and offer mobile banking, in the shortest possible time.

The proliferation of the 3G (third generation of wireless) and widespread implementation expected for 2003-2007 will generate the development of more sophisticated services such as multimedia and links to m-commerce services.

Opportunities In Mobile Banking

Account Information

1. Mini-statements and Checking account history
2. Term deposits
3. Loans statement
4. Cards statement
5. Mutual Funds / Equity statement
6. Insurance policy
7. Pension plan

Payments & Transfers

1. Domestic and international fund transfers
2. Micro-payments
3. Mobile re-charge
4. Commercial payments
5. Bill payment

Investments

1. Portfolio Management Services
2. Real-time stock quotes
3. Personalized alerts and notifications on security prices

Support

1. Status of origination of Mortgage, Insurance
2. Check book request
3. Exchange data message and email / Complaints

Content Services

1. General information such as Weather updates, News
2. Loyalty related offers
3. Location dependent services

Based on survey conducted by Forrester, mobile banking will be attractive mainly to the younger, more tech-savvy customer segment. A third of mobile phone users say that they may consider performing some kind of financial transaction through their mobile phone. But most of the users are interested in performing basic transactions such as querying for account balance and making bill payment.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

Challenges For A Mobile Banking Solution

Interoperability

There is a lack of common technology standards for mobile banking. Many protocols are being used for mobile banking – HTML, WAP, SOAP, XML to name a few. It would be a wise idea for the vendor to develop a mobile banking application that can connect multiple banks. It would require either the application to support multiple protocols or use of a common and widely acceptable set of protocols for data exchange.

There are a large number of different mobile phone devices and it is a big challenge for banks to offer mobile banking solution on any type of device. Some of these devices support J2ME and others support WAP browser or only SMS.

Security

Security of financial transaction, being executed from some remote location and transmission of financial information over the air, are the most complicated challenges that need to be addressed jointly by mobile application developers, wireless network service providers and the bank's IT department.

The following aspects need to be addressed to offer a secure infrastructure for financial transaction over wireless network:

1. Physical security of the handheld device. If the bank is offering smart-card based security, the physical security of the device is more important.
2. Security of the thick-client application running on the device. In case the device is stolen, the hacker should require ID/Password to access the application.
3. Authentication of the device with service provider before initiating a transaction. This would ensure that unauthorized devices are not connected to perform financial transactions.
4. User ID / Password authentication of bank's customer.
5. Encryption of the data being transmitted over the air.
6. Encryption of the data that will be stored in device for later / off-line analysis by the customer.

Scalability & Reliability

Another challenge for the CIOs and CTOs of the banks is to scale-up the mobile banking infrastructure to handle exponential growth of the customer base. With mobile banking, the customer may be sitting in any part of the world (a true anytime, anywhere banking) and hence banks need to ensure that the systems are up and running in a true 24 x 7 fashion. As customers will find mobile banking more and more useful, their expectations from the solution will increase. Banks unable to meet the performance and reliability expectations may lose customer confidence.

Application Distribution

Due to the nature of the connectivity between bank and its customers, it would be impractical to expect customers to regularly visit banks or connect to a web site for regular upgrade of their mobile banking application. It will be expected that the mobile application itself check the upgrades and updates and download necessary patches. However, there could be many issues to implement this approach such as upgrade / synchronization of other dependent components.

Personalization

It would be expected from the mobile application to support personalization such as :

1. Preferred Language
2. Date / Time format
3. Amount format
4. Default transactions
5. Standard Beneficiary list
6. Alerts

E-Commerce – Mobile Payments

What Is It?

Mobile payment is paying for goods or services with a mobile device such as a mobile phone, Personal Digital Assistant (PDA), or other such device. They can be used in a variety of payment scenarios. Typical usage entails the user electing to make a mobile payment, being connected to a server via the mobile device to perform authentication and authorization, and subsequently being presented with confirmation of the completed transaction.

Mobile Phone Micropayment

Mobile phone micropayment is a system of micropayment using mobile phones that is currently used predominantly in developing countries although the potential market is global.

4Sample Application: Parking

Mobile phone parking is a solution for payment of parking by use of a cellular phone. It is used both on on-street and off-street parking, but usually when a barrier is involved. Mobile Phone Parking may be said to substitue parking meters and pay and display machines. Parking wardens may enforce the parkings by license plate, transponder tags or barcode stickers.

Mobile Phone Parking offers both drivers and parking operators several distinct advantages. End users benefit from the convenience of being able to pay for parking from the comfort of their car with their mobile phone, and parking operators are not obliged to invest in either existing or new street-based parking infrastructures.

Mobile Phone Parking as a concept were established in the 1990, while commercialization started around 2000. A leading vendor in the business is the Scandinavian company Easy Park ASA.

Simpay

SimPay is a consortium which was founded to promote Mobile Payment but which is being closed as of June 24, 2005.

In February 2003, T-Mobile, Orange, and Vodafone formed a new Mobile Payment Services Association (MPSA) with the goal to deliver an open, interoperable and commonly branded solution for payments via mobile phones, designed to work across all operator networks. In June 2003 the consortium re-branded itself as SimPay.

In February 2005, Amena and Proximus joined the consortium. Several other European operators (Elisa, Mobilkom, Optimus, SFR, Telia, H3G, Debitel, KPN, O2) have also expressed an interest to join SimPay.

Simpay planned to create a pan-european framework whereby merchants and content resellers would be able to charge for products and services directly to a subscriber's bill.

In June 2005, Simpay decided, "following the decision of one of its founding Members not to launch Simpay for the foreseeable future, [...] not to pursue its activity on a pan-European scale as originally planned." All activities were put on hold effective June 24, 2005.

Organizations like Bango which were working with Simpay established multiple bi-lateral agreements and now provide similar services - across many more mobile operators.

The UK Simpay founders started work in 2005 on a project called "Payforit" which was launched in 2007 to provide similar services in the UK.

Mobile FeliCa

Mobile FeliCa has become the de-facto standard method for mobile payments in Japan, where mobile payments have become highly developed. Mobile payment means using your cell phone or other portable electronic device to pay for goods or services.

Mobile FeliCa is owned, developed and promoted by the company FeliCa Networks, which is a joint venture between Sony, NTT DoCoMo and East Japan Railway Company based on Sony's FeliCa contactless IC-Card technology.

Mobile Suica

Mobile Suica is a service for Osaifu Keitai mobile phones, first released on January 28, 2006 by NTT DoCoMo and au in Japan. Suica is a prepaid rechargeable contactless smart card mainly used to pay for fares on the JR East railway network.

Just like traditional Suica cards, Mobile Suica can also be charged when the remaining balance gets low. Other features supported by the mobile phone includes the ability to review past Suica transactions via the mobile's display. Suica uses Sony's FeliCa chip for its main functionalities. Mobile Suica interacts with the FeliCa chip using Java technology.

Since October 2006, it is possible to register for Mobile Suica using any major credit card, a Suica VIEW card is no longer required. A limited e-money only application called "Easy Mobile Suica" (which does not require a credit card) was also launched in late October 2006.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

E-Commerce – Smartcards

What Is It?

A smart card, chip card, or integrated circuit(s) card (ICC), is defined as any pocket-sized card with embedded integrated circuits. Although there is a diverse range of applications, there are two broad categories of ICCs. Memory cards contain only non-volatile memory storage components, and perhaps some specific security logic. Microprocessor cards contain memory and microprocessor components.

The standard perception of a "smart card" is a microprocessor card of credit card dimensions (or smaller, e.g. the mini GSM SIM card) with various tamper-resistant properties (e.g. a secure crypto-processor, secure file system, human-readable features) and is capable of providing security services (e.g. confidentiality of information in the memory). Not all chip cards contain a microprocessor (eg. the memory cards), therefore not all chip cards are necessarily also smart cards. However, the public usage of the terminology is often inconsistent.

History

Smart cards were invented and patented in the 1970s. There are some disputes regarding the actual "inventor"; claimants include Jürgen Dethloff of Germany, Kunitaka Arimura of Japan, and Roland Moreno of France. The first mass use of the cards was for payment in French pay phones, starting in 1983 (Télécarte).

Roland Moreno actually patented the concept of the memory card in 1974. In 1977, Michel Ugon from Honeywell Bull invented the first microprocessor smart card. In 1978, Bull patented the SPOM (Self Programmable One-chip Microcomputer) that defines the necessary architecture to auto-program the chip. Three years later, the very first "CP8" based on this patent was produced by Motorola. Today, Bull has 1200 patents related to smart cards.

The second use was with the integration of a microchips into all French debit cards (Carte Bleue) completed in 1992. When paying in France with a Carte Bleue, one inserts the card into the merchant's terminal, then types the PIN, before the transaction is accepted. Only very limited transactions (such as paying small autoroute tolls) are accepted without PIN.

Smart-card-based electronic purse systems (in which value is stored on the card chip, not in an externally recorded account, so that machines accepting the card need no network connectivity) were tried throughout Europe from the mid-1990s, most notably in Germany (Geldkarte), Austria (Quick), Belgium (Proton), the Netherlands (Chipknip and Chipper), Switzerland ("Cash"), Sweden ("Cash"), Finland ("Avant"), UK ("Mondex") and Denmark ("Danmønt").

The major boom in smart card use came in the 1990s, with the introduction of the smart-card-based SIM used in GSM mobile phone equipment in Europe. With the ubiquity of mobile phones in Europe, smart cards have become very common.

The international payment brands MasterCard, Visa, and Europay agreed in 1993 to work together to develop the specifications for the use of smart cards in payment cards used as either a debit or a credit card. The first version of the EMV system was released in 1994. In 1998 a stable release of the specifications was available. EMVco, the company responsible for the long-term maintenance of the system, upgraded the specification in 2000 and most recently in 2004. The goal of EMVco is to assure the various financial institutions and retailers that the specifications retain backward compatibility with the 1998 version.

With the exception of the United States there has been significant progress in the deployment of EMV-compliant point of sale equipment and the issuance of debit and or credit cards adhering the EMV specifications. Typically, a country's national payment association, in coordination with MasterCard International, Visa International, American Express and JCB, develop detailed implementation plans assuring a coordinated effort by the various stakeholders involved.

The backers of EMV claim it is a paradigm shift in the way one looks at payment systems. Though some banks are considering issuing one card that will serve as both a debit card and as a credit card, the business justification for this is still quite elusive. Within EMV a concept called Application Selection defines how the consumer selects which means of payment to employ for that purchase at the point of sale.

For the banks interested in introducing smart cards the only quantifiable benefit is the ability to forecast a significant reduction in fraud, in particular counterfeit, lost and stolen. The current level of fraud a country is experiencing determines if there is a business case for the financial institutions. Some critics claim that the savings are far less than the cost of implementing EMV, and thus many believe that the USA payments industry will opt to wait out the current EMV life cycle in order to implement new, contactless technology.

Smart cards with contactless interfaces are becoming increasingly popular for payment and ticketing applications such as mass transit. Visa and MasterCard have agreed to an easy-to-implement version currently being deployed (2004-2006) in the USA. Across the globe, contactless fare collection systems are being implemented to drive efficiencies in public transit. The various standards emerging are local in focus and are not compatible, though the MIFARE card from Phillips has a considerable market share in the US and Europe.

Smart cards are also being introduced in personal identification and entitlement schemes at regional, national, and international levels. Citizen cards, drivers' licences, and patient card schemes are becoming more prevalent, and contactless smart cards are being integrated into ICAO biometric passports to enhance security for international travel.

Contact Smart Card

Contact Smart Cards have a small gold chip about ½ inch in diameter on the front. When inserted into a reader, the chip makes contact with electrical connectors that can read information from the chip and write information back.

The ISO/IEC 7816 and ISO/IEC 7810 series of standards define:

- the physical shape
- the positions and shapes of the electrical connectors
- the electrical characteristics
- the communications protocols
- the format of the commands sent to the card and the responses returned by the card
- robustness of the card
- the functionality

The cards do not contain batteries; energy is supplied by the card reader.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

Contact Smart Card Reader

Contact smart card readers are used as a communications medium between the smart card and a host, e.g. a computer.

Contactless Smart Card

A second type is the contactless smart card, in which the chip communicates with the card reader through RFID induction technology (at data rates of 106 to 848 kbit/s). These cards require only close proximity to an antenna to complete transaction. They are often used when transactions must be processed quickly or hands-free, such as on mass transit systems, where smart cards can be used without even removing them from a wallet.

The standard for contactless smart card communications is ISO/IEC 14443, dated 2001. It defines two types of contactless cards ("A" and "B"), allows for communications at distances up to 10 cm. There had been proposals for ISO 14443 types C, D, E and F that have been rejected by the International Organization for Standardization. An alternative standard for contactless smart cards is ISO 15693, which allows communications at distances up to 50 cm.

Example of widely used contactless smart cards are Hong Kong's Octopus card, Paris' Calypso/Navigo card and Lisbon' LisboaViva card, which predate the ISO/IEC 14443 standard. The following tables list smart cards used for public transportation and other electronic purse applications.

A related contactless technology is RFID (radio frequency identification). In certain cases, it can be used for applications similar to those of contactless smart cards, such as for electronic toll collection. RFID devices usually do not include writeable memory or microcontroller processing capability as contactless smart cards often do.

There are dual-interface cards that implement contactless and contact interfaces on a single card with some shared storage and processing. An example is Porto's multi-application transport card, called Andante, that uses a chip in contact and contactless (ISO 14443B).

Like smart cards with contacts, contactless cards do not have a battery. Instead, they use a built-in inductor to capture some of the incident radio-frequency interrogation signal, rectify it, and use it to power the card's electronics.

Cryptographic Smart Cards

Most advanced smart cards are equipped with specialized cryptographic hardware that let you use algorithms such as RSA and DSA on board. Today's cryptographic smart cards are also able to generate key pairs on board, to avoid the risk of having more than one copy of the key (since by design (usually) there isn't a way to extract the keys from a smart card).

Such a smart cards are mainly used for digital signature and secure identification (see applications section).

The most common way to access to a cryptographic smart card functions on a computer is to use a PKCS#11 library provided by the vendor. On Microsoft Windows platforms the CSP API is also adopted.

Applications

Financial

The applications of smart cards include their use as credit or ATM cards, in a fuel card, SIMs for mobile phones, authorization cards for pay television, high-security identification and access-control cards, and public transport and public phone payment cards.

Smart cards may also be used as electronic wallets. The smart card chip can be loaded with funds which can be spent in parking meters and vending machines or at various merchants. Cryptographic protocols protect the exchange of money between the smart card and the accepting machine.

Identification

A quickly growing application is in digital identification cards. In this application, the cards are used for authentication of identity. The most common example is in conjunction with a PKI. The smart card will store an encrypted digital certificate issued from the PKI along with any other relevant or needed information about the card holder. Examples include the U.S. Department of Defense (DoD) Common Access Card (CAC), and the use of various smart cards by many governments as identification cards for their citizens. When combined with biometrics, smart cards can provide two- or three-factor authentication. Smart cards are a privacy-enhancing technology, for the subject carries possibly incriminating information about him all the time. By employing contactless smart cards, that can be read without having to remove the card from the wallet or even the garment it is in, one can add even more authentication value to the human carrier of the cards.

The first smart card driver's license system in the world was issued in 1995 in Mendoza, a province of Argentina. Mendoza has a high level of road accidents, driving offenses, and a poor record of recovering outstanding fines.[citation needed] The smart licenses keep an up-to-date record of driving offenses and unpaid fines. They also store personal information, license type and number, and a photograph of the holder. Emergency medical information like blood type, allergies, and biometrics (fingerprints) can be stored on the chip if the cardholder wishes. The Argentina government anticipates that this new system will help to recover more than $10 million per year in fines.

Gujarat was the first state in India to introduce the smart card license system in 1999. To date the Gujarat Government has issued 5 million smart card driving licenses to its people.[citation needed] This card is basically a plastic card having ISO 7810 certification and integrated circuit, capable of storing and verifying information according to its programming.

Smart cards have been advertised as suitable for personal identification tasks, because they are engineered to be tamper resistant. The embedded chip of a smart card usually implements some cryptographic algorithm. Information about the inner workings of this algorithm can be obtained if the precise time and electrical current required for certain encryption or decryption operations is measured. A number of research projects have now demonstrated the feasibility of this line of attack. Countermeasures have been proposed.

Other

Smart cards are widely used to protect digital television streams. See television encryption for an overview, and VideoGuard for a specific example of how smartcard security worked (and was cracked).

Problems

Another problem of smart cards may be the failure rate. The plastic card in which the chip is embedded is fairly flexible, and the larger the chip, the higher the probability of breaking. Smart cards are often carried in wallets or pockets — a fairly harsh environment for a chip. However, for large banking systems, the failure-management cost can be more than offset by the fraud reduction.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www*.planetlimited.*io

The Internet – The Internet

What Is It?

The Internet is the worldwide, publicly accessible network of interconnected computer networks that transmit data by packet switching using the standard Internet Protocol (IP). It is a "network of networks" that consists of millions of smaller domestic, academic, business, and government networks, which together carry various information and services, such as electronic mail, online chat, file transfer, and the interlinked Web pages and other documents of the World Wide Web.

The Internet As A "Series Of Tubes"

"Series of tubes" was a metaphor used by United States Senator and then-Chairman of the United States Senate Committee on Commerce, Science and Transportation Ted Stevens (R-Alaska). He was trying to describe the Internet in a speech about network neutrality.[1] Stevens was criticizing a proposed amendment to a committee bill which would have prohibited Internet service providers from charging fees to give some companies higher priority access to their networks or their customers. The metaphor became emblematic of the speech (and Stevens' seemingly poor understanding of the Internet) despite Stevens making several other odd comparisons and references. Stevens' use of the metaphor has been defended by experts in the field.

The speech took place on Wednesday, June 28, 2006.

Partial Text Of Stevens' Comments

"Ten movies streaming across that, that Internet, and what happens to your own personal Internet? I just the other day got... an Internet was sent by my staff at 10 o'clock in the morning on Friday, I got it yesterday. Why? Because it got tangled up with all these things going on the Internet commercially."

"They want to deliver vast amounts of information over the Internet. And again, the Internet is not something you just dump something on. It's not a big truck. It's a series of tubes. And if you don't understand those tubes can be filled and if they are filled, when you put your message in, it gets in line and it's going to be delayed by anyone that puts into that tube enormous amounts of material, enormous amounts of material."

Defense Of Terminology

Although this Author personally cringes at Steven's speech, it was notably defended by Princeton computer science professor Edward Felten, who said that he disagreed with Stevens' argument but felt that his terminology was entirely reasonable as a non-technical explanation given off-the-cuff in a meeting.

The term pipe is a commonly used idiom to refer to a data connection, with pipe size being analogous to bandwidth.

Routers use a data structure called a queue to buffer packets. When packets arrive more quickly than can be forwarded, the router will hold the packets in a queue until they can be sent on to the next router or be dropped. On congested links packets typically spend more time waiting in the queue than they do actually moving down wires or optical fiber. It is the delay of packets in the queue that causes the latency problems that make certain types of services impossible to use.

Terminology: Internet Versus Web

The Internet and the World Wide Web are not synonymous: the Internet is a collection of interconnected computer networks, linked by copper wires, fiber-optic cables, wireless connections, etc.; the Web is a collection of interconnected documents and other resources, linked by hyperlinks and URLs. The World Wide Web is accessible via the Internet, as are many other services including e-mail, file sharing, and others described below.

The best way to define and distinguish between these terms is with reference to the Internet protocol suite. This collection of standards and protocols is organized into layers such that each layer provides the foundation and the services required by the layer above. In this conception, the term Internet refers to computers and networks that communicate using IP (Internet protocol) and TCP (transfer control protocol). Once this networking structure is established, then other protocols can run “on top.” These other protocols are sometimes called services or applications. Hypertext transfer protocol, or HTTP, is the application layer protocol that links and provides access to the files, documents and other resources of the World Wide Web.

Creation Of The Internet

The USSR's launch of Sputnik spurred the United States to create the Advanced Research Projects Agency (ARPA, later known as the Defense Advanced Research Projects Agency, or DARPA) in February 1958 to regain a technological lead. ARPA created the Information Processing Technology Office (IPTO) to further the research of the Semi Automatic Ground Environment (SAGE) program, which had networked country-wide radar systems together for the first time. J. C. R. Licklider was selected to head the IPTO, and saw universal networking as a potential unifying human revolution.

In 1950, Licklider moved from the Psycho-Acoustic Laboratory at Harvard University to MIT where he served on a committee that established MIT Lincoln Laboratory. He worked on the SAGE project. In 1957 he became a Vice President at BBN, where he bought the first production PDP-1 computer and conducted the first public demonstration of time-sharing.

Licklider recruited Lawrence Roberts to head a project to implement a network, and Roberts based the technology on the work of Paul Baran who had written an exhaustive study for the U.S. Air Force that recommended packet switching (as opposed to Circuit switching) to make a network highly robust and survivable. After much work, the first node went live at UCLA on October 29, 1969 on what would be called the ARPANET, one of the "eve" networks of today's Internet. Following on from this, the British Post Office, Western Union International and Tymnet collaborated to create the first international packet switched network, referred to as the International Packet Switched Service (IPSS), in 1978. This network grew from Europe and the US to cover Canada, Hong Kong and Australia by 1981.

The first TCP/IP wide area network was operational by 1 January 1983, when the United States' National Science Foundation (NSF) constructed a university network backbone that would later become the NSFNet. (This date is held by some to be technically that of the birth of the Internet.) It was then followed by the opening of the network to commercial interests in 1985. Important, separate networks that offered gateways into, then later merged with, the NSFNet include Usenet, BITNET and the various commercial and educational X.25

Compuserve and JANET. Telenet (later called Sprintnet), was a large privately-funded national computer network with free dialup access in cities throughout the U.S. that had been in operation since the 1970s. This network eventually merged with the others in the 1990s as the TCP/IP protocol became increasingly popular. The ability of TCP/IP to work over these pre-existing communication networks, especially the international X.25 IPSS network, allowed for a great ease of growth. Use of the term "Internet" to describe a single global TCP/IP network originated around this time.

The network gained a public face in the 1990s. On August 6, 1991 CERN, which straddles the border between France and Switzerland publicized the new World Wide Web project, two years after Tim Berners-Lee had begun creating HTML, HTTP and the first few Web pages at CERN.

An early popular Web browser was ViolaWWW based upon HyperCard. It was eventually replaced in popularity by the Mosaic Web Browser. In 1993 the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign released version 1.0 of Mosaic and by late 1994 there was growing public interest in the previously academic/technical Internet. By 1996 the word "Internet" was coming into common daily usage, frequently misused to refer to the World Wide Web.

Meanwhile, over the course of the decade, the Internet successfully accommodated the majority of previously existing public computer networks (although some networks such as FidoNet have remained separate). This growth is often attributed to the lack of central administration, which allows organic growth of the network, as well as the non-proprietary open nature of the Internet protocols, which encourages vendor interoperability and prevents any one company from exerting too much control over the network.

Today's Internet

Aside from the complex physical connections that make up its infrastructure, the Internet is facilitated by bi- or multi-lateral commercial contracts (e.g., peering agreements), and by technical specifications or protocols that describe how to exchange data over the network. Indeed, the Internet is essentially defined by its interconnections and routing policies.

As of March 10, 2007, 1.114 billion people use the Internet according to Internet World Stats.

Internet Protocols

In this context, there are three layers of protocols:

- At the lowest level is IP (Internet Protocol), which defines the datagrams or packets that carry blocks of data from one node to another. The vast majority of today's Internet uses version four of the IP protocol (i.e. IPv4), and although IPv6 is standardized, it exists only as "islands" of connectivity, and there are many ISPs who don't have any IPv6 connectivity at all.
- Next come TCP (Transmission Control Protocol) and UDP (User Datagram Protocol) - the protocols by which one host sends data to another. The former makes a virtual 'connection', which gives some level of guarantee of reliability. The latter is a best-effort, connectionless transport, in which data packets that are lost in transit will not be re-sent.
- On top comes the application protocol. This defines the specific messages and data formats sent and understood by the applications running at each end of the communication.

Internet Structure

There have been many analyses of the Internet and its structure. For example, it has been determined that the Internet IP routing structure and hypertext links of the World Wide Web are examples of scale-free networks.

Similar to the way the commercial Internet providers connect via Internet exchange points, research networks tend to interconnect into large sub-networks such as:

- GEANT
- GLORIAD
- Abilene Network
- JANET (the UK's Joint Academic Network aka UKERNA)

These in turn are built around relatively smaller networks. See also the list of academic computer network organizations

In network schematic diagrams, the Internet is often represented by a cloud symbol, into and out of which network communications can pass.

ICANN

The Internet Corporation for Assigned Names and Numbers (ICANN) is the authority that coordinates the assignment of unique identifiers on the Internet, including domain names, Internet Protocol (IP) addresses, and protocol port and parameter numbers. A globally unified namespace (i.e., a system of names in which there is one and only one holder of each name) is essential for the Internet to function. ICANN is headquartered in Marina del Rey, California, but is overseen by an international board of directors drawn from across the Internet technical, business, academic, and non-commercial communities. The US government continues to have the primary role in approving changes to the root zone file that lies at the heart of the domain name system. Because the Internet is a distributed network comprising many voluntarily interconnected networks, the Internet, as such, has no governing body. ICANN's role in coordinating the assignment of unique identifiers distinguishes it as perhaps the only central coordinating body on the global Internet, but the scope of its authority extends only to the Internet's systems of domain names, IP addresses, and protocol port and parameter numbers.

On Nov. 16, 2005, the World Summit on the Information Society, held in Tunis, established the Internet Governance Forum (IGF) to discuss Internet-related issues.

Language

The prevalent language for communication on the Internet is English. This may be a result of the Internet's origins, as well as English's role as the lingua franca. It may also be related to the poor capability of early computers to handle characters other than those in the basic Latin alphabet.

After English (30% of Web visitors) the most-requested languages on the World Wide Web are Chinese 14%, Japanese 8%, Spanish 8%, German 5%, and French 5% (from Internet World Stats, updated January 11, 2007).

By continent, 36% of the world's Internet users are based in Asia, 29% in Europe, and 21% in North America ([2] updated January 11, 2007).

The Internet's technologies have developed enough in recent years that good facilities are available for development and communication in most widely used languages. However, some glitches such as mojibake (incorrect display of foreign language characters, also known as krakozyabry) still remain.

Internet & The Workplace

The Internet is allowing greater flexibility in working hours and location, especially with the spread of unmetered high-speed connections and Web applications.

The Mobile Internet

The Internet can now be accessed virtually anywhere by numerous means. Mobile phones, datacards, and cellular routers allow users to connect to the Internet from anywhere there is a cellular network supporting that device's technology.

Common Uses Of The Internet

E-Mail

The concept of sending electronic text messages between parties in a way analogous to mailing letters or memos predates the creation of the Internet. Even today it can be important to distinguish between Internet and internal e-mail systems. Internet e-mail may travel and be stored unencrypted on many other machines and networks out of both the sender's and the recipient's control. During this time it is quite possible for the content to be read and even tampered with by third parties, if anyone considers it important enough. Purely internal or intranet mail systems, where the information never leaves the corporate or organization's network and servers, is much more secure, although in any organization there will be IT and other personnel whose job may involve monitoring, or at least occasionally accessing, the email of other employees not addressed to them. Web-based email (webmail) between parties on the same webmail system may not actually 'go' anywhere—it merely sits on the one server and is tagged in various ways so as to appear in one person's 'sent items' list and in one or more others' 'in boxes' or other 'folders' when viewed.

E-mail attachments have greatly increased the usefulness of e-mail in many ways. When a file is attached to an email, a text representation of the attached data (which may itself be binary data) is actually appended to the e-mail text, later to be reconstituted into a 'file' on the recipient's machine for their use. See MIME (Multipurpose Internet Mail Extensions) for details of how the problems involved in doing this have been overcome.

The World Wide Web

Through keyword-driven Internet research using search engines, like Google, millions worldwide have easy, instant access to a vast and diverse amount of online information. Compared to encyclopedias and traditional libraries, the World Wide Web has enabled a sudden and extreme decentralization of information and data.

Many individuals and some companies and groups have adopted the use of "Web logs" or blogs, which are largely used as easily-updatable online diaries. Some commercial organizations encourage staff to fill them with advice on their areas of specialization in the hope that visitors will be impressed by the expert knowledge and free information, and be attracted to the corporation as a result. One example of this practice is Microsoft, whose product developers publish their personal blogs in order to pique the public's interest in their work.

For more information on the distinction between the World Wide Web and the Internet itself — as in everyday use the two are sometimes confused — see Dark internet where this is discussed in more detail.

Remote Access

The Internet allows computer users to connect to other computers and information stores easily, wherever they may be across the world. They may do this with or without the use of security, authentication and encryption technologies, depending on the requirements.

This is encouraging new ways of working from home, collaboration and information sharing in many industries. An accountant sitting at home can audit the books of a company based in another country, on a server situated in a third country that is remotely maintained by IT specialists in a fourth. These accounts could have been created by home-working book-keepers, in other remote locations, based on information e-mailed to them from offices all over the world. Some of these things were possible before the widespread use of the Internet, but the cost of private, leased lines would have made many of them infeasible in practice.

An office worker away from his desk, perhaps the other side of the world on a business trip or a holiday, can open a remote desktop session into his normal office PC using a secure Virtual Private Network (VPN) connection via the Internet. This gives him complete access to all his normal files and data, including e-mail and other applications, while he is away.

This concept is also referred to by some network security people as the Virtual Private Nightmare, because it extends the secure perimeter of a corporate network into its employees' homes; this has been the source of some notable security breaches, but also provides security for the workers.

Collaboration

The low-cost and nearly instantaneous sharing of ideas, knowledge, and skills has made collaborative work dramatically easier. Not only can a group cheaply communicate and test, but the wide reach of the Internet allows such groups to easily form in the first place, even among niche interests. An example of this is the free software movement in software development which produced GNU and Linux from scratch and has taken over development of Mozilla and OpenOffice.org (formerly known as Netscape Communicator and StarOffice).

Internet 'chat', whether in the form of IRC 'chat rooms' or channels, or via instant messaging systems allow colleagues to stay in touch in a very convenient way when working at their computers during the day. Messages can be sent and viewed even more quickly and conveniently than via e-mail. Extension to these systems may allow files to be exchanged, 'whiteboard' drawings to be shared as well as voice and video contact between team members.

Version control systems allow collaborating teams to work on shared sets of documents without either accidentally overwriting each other's work or having members wait until they get 'sent' documents to be able to add their thoughts and changes.

File Sharing

A computer file can be e-mailed to customers, colleagues and friends as an attachment. It can be uploaded to a Web site or FTP server for easy download by others. It can be put into a "shared location" or onto a file server for instant use by colleagues. The load of bulk downloads to many users can be eased by the use of "mirror" servers or peer-to-peer networks. In any of these cases, access to the file may be controlled by user authentication; the transit of the file over the Internet may be obscured by encryption and money may change hands before or after access to the file is given. The price can be paid by the remote charging of funds from, for example a credit card whose details are also passed - hopefully fully encrypted - across the Internet. The origin and authenticity of the file received may be checked by digital signatures or by MD5 or other message digests.

These simple features of the Internet, over a world-wide basis, are changing the basis for the production, sale, and distribution of anything that can be reduced to a computer file for transmission. This includes all manner of office documents, publications, software products, music, photography, video, animations, graphics and the other arts. This in turn is causing seismic shifts in each of the existing industry associations, such as the RIAA and MPAA in the United States, that previously controlled the production and distribution of these products in that country.

Streaming Media

Many existing radio and television broadcasters provide Internet 'feeds' of their live audio and video streams (for example, the BBC). They may also allow time-shift viewing or listening such as Preview, Classic Clips and Listen Again features. These providers have been joined by a range of pure Internet 'broadcasters' who never had on-air licenses. This means that an Internet-connected device, such as a computer or something more specific, can be used to access on-line media in much the same way as was previously possible only with a TV or radio receiver. The range of material is much wider, from pornography to highly specialized technical Web-casts. Podcasting is a variation on this theme, where—usually audio—material is first downloaded in full and then may be played back on a computer or shifted to a digital audio player to be listened to on the move. These techniques using simple equipment allow anybody, with little censorship or licensing control, to broadcast audio-visual material on a worldwide basis.

Webcams can be seen as an even lower-budget extension of this phenomenon. While some webcams can give full frame rate video, the picture is usually either small or updates slowly. Internet users can watch animals around an African waterhole, ships in the Panama Canal, the traffic at a local roundabout or their own premises, live and in real time. Video chat rooms, video conferencing, and remote controllable webcams are also popular. Many uses can be found for personal webcams in and around the home, with and without two-way sound.

Voice Telephony (VoIP)

VoIP stands for Voice over IP, where IP refers to the Internet Protocol that underlies all Internet communication. This phenomenon began as an optional two-way voice extension to some of the Instant Messaging systems that took off around the year 2000. In recent years many VoIP systems have become as easy to use and as convenient as a normal telephone. The benefit is that, as the Internet carries the actual voice traffic, VoIP can be free or cost much less than a normal telephone call, especially over long distances and especially for those with always-on ADSL or DSL Internet connections.

Thus VoIP is maturing into a viable alternative to traditional telephones. Interoperability between different providers has improved and the ability to call or receive a call from a traditional telephone is available. Simple inexpensive VoIP modems are now available that eliminate the need for a PC.

Voice quality can still vary from call to call but is often equal to and can even exceed that of traditional calls.

Remaining problems for VoIP include emergency telephone number dialing and reliability. Currently a few VoIP providers provide some 911 dialing but it is not universally available. Traditional phones are line powered and operate during a power failure, VoIP does not do so without a backup power source for the electronics.

Most VoIP providers offer unlimited national calling but the direction in VoIP is clearly toward global coverage with unlimited minutes for a low monthly fee.

VoIP has also become increasingly popular within the gaming world, as a form of communication between players. Popular gaming VoIP clients include Ventrilo and Teamspeak, and there are others available also.

Censorship

Some governments, such as those of Iran, the People's Republic of China and Cuba restrict what people in their countries can access on the Internet, especially political and religious content. This is accomplished through software that filters domains and content so that they may not be easily accessed or obtained without elaborate circumvention.

In Norway, Finland and Sweden, major Internet service providers have voluntarily (possibly to avoid such an arrangement being turned into law) agreed to restrict access to sites listed by police. While this list of forbidden URL's is only supposed to contain addresses of known child pornography sites, content of the list is secret.

Many countries have enacted laws making the possession or distribution of certain material, such as child pornography, illegal, but do not use filtering software.

There are many free and commercially available software programs with which a user can choose to block offensive Web sites on individual computers or networks, such as to limit a child's access to pornography or violence.

Internet Access

Public places to use the Internet include libraries and Internet cafes, where computers with Internet connections are available. There are also Internet access points in many public places such as airport halls and coffee shops, in some cases just for brief use while standing. Various terms are used, such as "public Internet kiosk", "public access terminal", and "Web payphone". Many hotels now also have public terminals, though these are usually fee based.

Wi-Fi provides wireless access to computer networks, and therefore can do so to the Internet itself. Hotspots providing such access include Wi-Fi-cafes, where a would-be user needs to bring their own wireless-enabled devices such as a laptop or PDA. These services may be free to all, free to customers only, or fee-based. A hotspot need not be limited to a confined location. The whole campus or park, or even the entire city can be enabled. Grassroots efforts have led to wireless community networks. Commercial WiFi services covering large city areas are in place in London, Vienna, San Francisco, Philadelphia, Chicago, Pittsburgh and other cities, including Toronto by the end of 2006. The Internet can then be accessed from such places as a park bench.

Apart from Wi-Fi, there have been experiments with proprietary mobile wireless networks like Ricochet, various high-speed data services over cellular phone networks, and fixed wireless services.

High-end mobile phones such as smartphones generally come with Internet access through the phone network. Web browsers such as Opera are available on these advanced handsets, which can also run a wide variety of other Internet software. More mobile phones have Internet access than PCs, though this is not as widely used. An internet access provider and protocol matrix differentiates the methods used to get online.

Leisure

The Internet has been a major source of leisure since before the World Wide Web, with entertaining social experiments such as MUDs and MOOs being conducted on university servers, and humour-related Usenet groups receiving much of the main traffic. Today, many Internet forums have sections devoted to games and funny videos; short cartoons in the form of Flash movies are also popular. Over 6 million people use blogs or message boards as a means of communication and for the sharing of ideas.

The pornography and gambling industries have both taken full advantage of the World Wide Web, and often provide a significant source of advertising revenue for other Web sites. Although many governments have attempted to put restrictions on both industries' use of the Internet, this has generally failed to stop their widespread popularity. A song in the Broadway musical show Avenue Q is titled "The Internet is for Porn" and refers to the popularity of this aspect of the internet.

One main area of leisure on the Internet is multiplayer gaming. This form of leisure creates communities, bringing people of all ages and origins to enjoy the fast-paced world of multiplayer games. These range from MMORPG to first-person shooters, from role-playing games to online gambling. This has revolutionized the way many people interact and spend their free time on the Internet.

While online gaming has been around since the 1970s, modern modes of online gaming began with services such as GameSpy and MPlayer, which players of games would typically subscribe to. Non-subscribers were limited to certain types of gameplay or certain games.

Many use the Internet to access and download music, movies and other works for their enjoyment and relaxation. As discussed above, there are paid and unpaid sources for all of these, using centralized servers and distributed peer-to-peer technologies. Discretion is needed as some of these sources take more care over the original artists' rights and over copyright laws than others.

Many use the World Wide Web to access news, weather and sports reports, to plan and book holidays and to find out more about their random ideas and casual interests.

People use chat, messaging and email to make and stay in touch with friends worldwide, sometimes in the same way as some previously had pen pals. Social networking Web sites like Friends Reunited and many others like them also put and keep people in contact for their enjoyment.

The Internet has seen a growing amount of Internet operating systems, where users can access their files, folders, and settings via the Internet. An example of an opensource webos is Eyeos.

Cyberslacking has become a serious drain on corporate resources; the average UK employee spends 57 minutes a day surfing the Web at work, according to a study by Peninsula Business Services.

Complex Architecture

Many computer scientists see the Internet as a "prime example of a large-scale, highly engineered, yet highly complex system". The Internet is extremely heterogeneous. (For instance, data transfer rates and physical characteristics of connections vary widely.) The Internet exhibits "emergent phenomena" that depend on its large-scale organization. For example, data transfer rates exhibit temporal self-similarity. Further adding to the complexity of the Internet is the ability of more than one computer to use the Internet through only one node, thus creating the possibility for a very deep and hierarchal based sub-network that can theoretically be extended infinitely (disregarding the programmatic limitations of the IPv4 protocol).

Marketing

The Internet has also become a large market for companies; some of the biggest companies today have grown by taking advantage of the efficient nature of low-cost advertising and commerce through the Internet; also known as e-commerce. It is the fastest way to spread information to a vast amount of people simultaneously. The Internet has also subsequently revolutionized shopping—for example; a person can order a CD online and receive it in the mail within a couple of days, or download it directly in some cases. The Internet has also greatly facilitated personalized marketing which allows a company to market a product to a specific person or a specific group of people more so than any other advertising medium.

Examples of personalized marketing include online communities such as MySpace, Friendster, Orkut, and others which thousands of Internet users join to advertise themselves and make friends online. Many of these users are young teens and adolescents ranging from 13 to 25 years old. In turn, when they advertise themselves they advertise interests and hobbies, which online marketing companies can use as information as to what those users will purchase online, and advertise their own companies' products to those users.

A very ineffective way of advertising on the Internet is through spamming an email with advertisements. This is ineffective because, now, most email providers offer protection against email spam. Most spam messages are sent automatically to everybody in the email database of the company/person that is spamming. This way of advertising is almost like using adware.

Adware is another ineffective way of advertising because most people simply close a popup window when it shows up, not bothering to read it.

The Name "Internet"

Internet is traditionally written with a capital first letter, as it is a proper noun. The Internet Society, the Internet Engineering Task Force, the Internet Corporation for Assigned Names and Numbers, the World Wide Web Consortium, and several other Internet-related organizations use this convention in their publications.

Many newspapers, newswires, periodicals, and technical journals capitalize the term (Internet). Examples include the New York Times, the Associated Press, Time, The Times of India, Hindustan Times, and Communications of the ACM.

Others assert that the first letter should be written in lower case (internet). A significant number of publications use this form, including The Economist, the Canadian Broadcasting Corporation, the Financial Times, The Guardian, The Times, and The Sydney Morning Herald. As of 2005, many publications using internet appear to be located outside of North America—although one U.S. news source, Wired News, has adopted the lower case spelling.

Historically, Internet and internet have had different meanings, with internet being a contraction of internetwork or internetworking and Internet referring to a matrix of networks using TCP/IP (Transmission Controll Protocol/ Internet Protocol) according to the book Where wizards stay up late. Under this distinction, the Internet is a particular internet, but the reverse does not apply. The distinction was evident in many RFCs, books, and articles from the 1980s and early 1990s (some of which, such as RFC 1918, refer to "internets" in the plural), but has recently fallen into disuse.[citation needed] Instead, the term intranet is generally used for private networks.

Some people use the lower-case term as a medium (like radio or newspaper, e.g. I've found it in internet), and capitalised (or first letter capitalised) as the global network.

Significant Internet Events

Malfunctions & Attacks

- Morris worm — November 2, 1988
- Predicted Y2K Bug - January 1, 2000
- UUNet/Worldcom backbone difficulties — October 3, 2002
- 2002 DNS Backbone DDoS — October 22, 2002
- SQL Slammer worm — January 24, 2003

The Internet – The World Wide Web

What Is It?

The World Wide Web ("WWW" or simply the "Web") is a system of interlinked, hypertext documents that runs over the Internet. With a Web browser, a user views Web pages that may contain text, images, and other multimedia and navigates between them using hyperlinks. The Web was created around 1990 by Tim Berners-Lee and the Belgian Robert Cailliau working at CERN in Geneva, Switzerland. Since then, Berners-Lee has played an active role in guiding the development of Web standards (such as the markup languages in which Web pages are composed), and in recent years has advocated his vision of a Semantic Web.

Basic Terms

The World Wide Web is the combination of four basic ideas:

- Hypertext: a format of information which allows, in a computer environment, one to move from one part of a document to another or from one document to another through internal connections among these documents (called "hyperlinks");
- Resource Identifiers: unique identifiers used to locate a particular resource (computer file, document or other resource) on the network - this is commonly known as a URL or URI, although the two have subtle technical differences;
- The Client-server model of computing: a system in which client software or a client computer makes requests of server software or a server computer that provides the client with resources or services, such as data or files; and
- Markup language: characters or codes embedded in text which indicate structure, semantic meaning, or advice on presentation.

On the World Wide Web, a client program called a user agent retrieves information resources, such as Web pages and other computer files, from Web servers using their URLs. If the user agent is a kind of Web browser, it displays the resources on a user's computer. The user can then follow hyperlinks in each web page to other World Wide Web resources, whose location is embedded in the hyperlinks. It is also possible, for example by filling in and submitting web forms, to post information back to a Web server for it to save or process in some way. Web pages are often arranged in collections of related material called "Web sites." The act of following hyperlinks from one Web site to another is referred to as "browsing" or sometimes as "surfing" the Web.

The phrase "surfing the Internet" was first made popular in print by Jean Armour Polly, a librarian, in an article called "Surfing the INTERNET", published in the University of Minnesota Wilson Library Bulletin in June 1992. Although Polly may have developed the phrase independently, slightly earlier uses of similar terms appeared on Usenet in 1991 and 1992, and some recollections claim it was also used verbally in the hacker community for a couple years before that.

For more information on the distinction between the World Wide Web and the Internet itself— as in everyday use the two are sometimes confused.

Although the English word worldwide is normally written as one word (without a space or hyphen), the proper name World Wide Web and abbreviation WWW are now well-established even in formal English. The earliest references to the Web called it the WorldWideWeb (an example of computer programmers' fondness for CamelCase) or the World-Wide Web (with a hyphen, this version of the name is the closest to normal English usage).

Ironically, the abbreviation "WWW" is somewhat impractical as it contains two or three times as many syllables (depending on accent) as the full term "World Wide Web", and thus takes longer to say.

How The Web Works

Viewing a Web page or other resource on the World Wide Web normally begins either by typing the URL of the page into a Web browser, or by following a hypertext link to that page or resource. The first step, behind the scenes, is for the server-name part of the URL to be resolved into an IP address by the global, distributed Internet database known as the Domain name system or DNS. The browser then establishes a TCP connection with the server at that IP address.

The next step is for an HTTP request to be sent to the Web server, requesting the resource. In the case of a typical Web page, the HTML text is first requested and parsed by the browser, which then makes additional requests for graphics and any other files that form a part of the page in quick succession. When considering web site popularity statistics, these additional file requests give rise to the difference between one single 'page view' and an associated number of server 'hits'.

The Web browser then renders the page as described by the HTML, CSS and other files received, incorporating the images and other resources as necessary. This produces the on-screen page that the viewer sees.

Most Web pages will themselves contain hyperlinks to other related pages and perhaps to downloads, source documents, definitions and other Web resources.

Such a collection of useful, related resources, interconnected via hypertext links, is what has been dubbed a 'web' of information. Making it available on the Internet created what Tim Berners-Lee first called the WorldWideWeb (note the name's use of CamelCase, subsequently discarded) in 1990.

Caching

If the user returns to a page fairly soon, it is likely that the data will not be retrieved from the source Web server, as above, again. By default, browsers cache all web resources on the local hard drive. An HTTP request will be sent by the browser that asks for the data only if it has been updated since the last download. If it has not, the cached version will be reused in the rendering step.

This is particularly valuable in reducing the amount of Web traffic on the Internet. The decision about expiration is made independently for each resource (image, stylesheet, JavaScript file etc., as well as for the HTML itself). Thus even on sites with highly dynamic content, many of the basic resources are only supplied once per session or less. It is worth it for any Web site designer to collect all the CSS and JavaScript into a few site-wide files so that they can be downloaded into users' caches and reduce page download times and demands on the server.

There are other components of the Internet that can cache Web content. The most common in practice are often built into corporate and academic firewalls where they cache web resources requested by one user for the benefit of all. Some search engines such as Google or Yahoo! also store cached content from Web sites.

Apart from the facilities built into Web servers that can ascertain when physical files have been updated, it is possible for designers of dynamically generated web pages to control the HTTP headers sent back to requesting users, so that pages are not cached when they should not be — for example Internet banking and news pages.

This helps with understanding the difference between the HTTP 'GET' and 'POST' verbs - data requested with a GET may be cached, if other conditions are met, whereas data obtained after POSTing information to the server usually will not.

History

The underlying ideas of the Web can be traced as far back as 1980, when, at CERN in Switzerland, the Englishman Tim Berners-Lee built ENQUIRE (referring to Enquire Within Upon Everything, a book he recalled from his youth). While it was rather different from the Web in use today, it contained many of the same core ideas (and even some of the ideas of Berners-Lee's next project after the WWW, the Semantic Web).

In March 1989, Tim Berners-Lee wrote Information Management: A Proposal, which referenced ENQUIRE and described a more elaborate information management system. With help from Robert Cailliau, he published a more formal proposal for the World Wide Web on November 12, 1990.

A NeXTcube was used by Berners-Lee as the world's first web server and also to write the first web browser, WorldWideWeb in 1990. By Christmas 1990, Berners-Lee had built all the tools necessary for a working Web [1]: the first Web browser (which was a Web editor as well), the first Web server and the first Web pages which described the project itself.

On August 6, 1991, he posted a short summary of the World Wide Web project on the alt.hypertext newsgroup. This date also marked the debut of the Web as a publicly available service on the Internet.

The crucial underlying concept of hypertext originated with older projects from the 1960s, such as Ted Nelson's Project Xanadu and Douglas Engelbart's oN-Line System (NLS). Both Nelson and Engelbart were in turn inspired by Vannevar Bush's microfilm-based "memex," which was described in the 1945 essay "As We May Think".

Berners-Lee's breakthrough was to marry hypertext to the Internet. In his book Weaving The Web, he explains that he had repeatedly suggested that a marriage between the two technologies was possible to members of both technical communities, but when no one took up his invitation, he finally tackled the project himself. In the process, he developed a system of globally unique identifiers for resources on the Web and elsewhere: the Uniform Resource Identifier.

The World Wide Web had a number of differences from other hypertext systems that were then available:

- The WWW required only unidirectional links rather than bidirectional ones. This made it possible for someone to link to another resource without action by the owner of that

resource. It also significantly reduced the difficulty of implementing Web servers and browsers (in comparison to earlier systems), but in turn presented the chronic problem of link rot.
- Unlike predecessors such as HyperCard, the World Wide Web was non-proprietary, making it possible to develop servers and clients independently and to add extensions without licensing restrictions.

On April 30, 1993, CERN announced that the World Wide Web would be free to anyone, with no fees due. Coming two months after the announcement that gopher was no longer free to use, this produced a rapid shift away from gopher and towards the Web. An early popular Web browser was ViolaWWW which was based upon HyperCard.

Scholars generally agree, however, that the turning point for the World Wide Web began with the introduction of the Mosaic web browser in 1993, a graphical browser developed by a team at the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign (NCSA-UIUC), led by Marc Andreessen. Funding for Mosaic came from the High-Performance Computing and Communications Initiative, a funding program initiated by then-Senator Al Gore's High Performance Computing and Communication Act of 1991 also known as the Gore Bill. Prior to the release of Mosaic, graphics were not commonly mixed with text in Web pages and its popularity was less than older protocols in use over the Internet, such as Gopher protocol and Wide area information server. Mosaic's graphical user interface allowed the Web to become by far the most popular Internet protocol.

Web Standards

At its core, the Web is made up of three standards:

- the Uniform Resource Identifier (URI), which is a universal system for referencing resources on the Web, such as Web pages;
- the HyperText Transfer Protocol (HTTP), which specifies how the browser and server communicate with each other; and
- the HyperText Markup Language (HTML), used to define the structure and content of hypertext documents.

Berners-Lee now heads the World Wide Web Consortium (W3C), which develops and maintains these and other standards that enable computers on the Web to effectively store and communicate different forms of information.

Java & JavaScript

A significant advance in Web technology was Sun Microsystems' Java platform. It enables Web pages to embed small programs (called applets) directly into the view. These applets run on the end-user's computer, providing a richer user interface than simple web pages. Java client-side applets never gained the popularity that Sun had hoped for, for a variety of reasons including lack of integration with other content (applets were confined to small boxes within the rendered page) and the fact that many computers at the time were supplied to end users without a suitably installed JVM, and so required a download by the user before applets would appear. Adobe Flash now performs many of the functions that were originally envisioned for Java applets including the playing of video content, animation and some rich UI features. Java itself has become more widely used as a platform and language for server-side and other programming.

JavaScript, on the other hand, is a scripting language that was initially developed for use within Web pages. The standardized version is ECMAScript. While its name is similar to Java, JavaScript was developed by Netscape and it has almost nothing to do with Java, apart from that, like Java, its syntax is derived from the C programming language. In conjunction with a Web page's Document Object Model, JavaScript has become a much more powerful technology than its creators originally envisioned. The manipulation of a page's Document Object Model after the page is delivered to the client has been called Dynamic HTML (DHTML), to emphasize a shift away from static HTML displays.

In its simplest form, all the optional information and actions available on a JavaScripted Web page will have been downloaded when the page was first delivered. Ajax ("Asynchronous JavaScript And XML") is a JavaScript-based technology that may have a significant effect on the development of the World Wide Web. Ajax provides a method whereby large or small parts within a Web page may be updated, using new information obtained over the network in response to user actions. This allows the page to be much more responsive, interactive and interesting, without the user having to wait for whole-page reloads. Ajax is seen as an important aspect of what is being called Web 2.0. Examples of Ajax techniques currently in use can be seen in Gmail, Google Maps etc.

Sociological Implications

The Web, as it stands today, has allowed global interpersonal exchange on a scale unprecedented in human history. People separated by vast distances, or even large amounts of time, can use the Web to exchange—or even mutually develop—their most intimate and extensive thoughts, or alternately their most casual attitudes and spirits. Emotional experiences, political ideas, cultural customs, musical idioms, business advice, artwork, photographs, literature, can all be shared and disseminated digitally with less individual investment than ever before in human history. Although the existence and use of the Web relies upon material technology, which comes with its own disadvantages, its information does not use physical resources in the way that libraries or the printing press have. Therefore, propagation of information via the Web (via the Internet, in turn) is not constrained by movement of physical volumes, or by manual or material copying of information. By virtue of being digital, the information of the Web can be searched more easily and efficiently than any library or physical volume, and vastly more quickly than a person could retrieve information about the world by way of physical travel or by way of mail, telephone, telegraph, or any other communicative medium.

The Web is the most far-reaching and extensive medium of personal exchange to appear on Earth. It has probably allowed many of its users to interact with many more groups of people, dispersed around the planet in time and space, than is possible when limited by physical contact or even when limited by every other existing medium of communication combined.

Because the Web is global in scale, some have suggested that it will nurture mutual understanding on a global scale. By definition or by necessity, the Web has such a massive potential for social exchange, it has the potential to nurture empathy and symbiosis, but it also has the potential to incite belligerence on a global scale, or even to empower demagogues and repressive regimes in ways that were historically impossible to achieve previously.

Publishing Web Pages

The Web is available to individuals outside mass media. In order to "publish" a Web page, one does not have to go through a publisher or other media institution, and potential readers could be found in all corners of the globe.

Unlike books and documents, hypertext does not need to have a linear order from beginning to end. It is not necessarily broken down into the hierarchy of chapters, sections, subsections, etc.

Many different kinds of information are now available on the Web, and for those who wish to know other societies, their cultures and peoples, it has become easier. When traveling in a foreign country or a remote town, one might be able to find some information about the place on the Web, especially if the place is in one of the developed countries. Local newspapers, government publications, and other materials are easier to access, and therefore the variety of information obtainable with the same effort may be said to have increased, for the users of the Internet.

Although some Web sites are available in multiple languages, many are in the local language only. Additionally, not all software supports all special characters, and RTL languages. These factors would challenge the notion that the World Wide Web will bring a unity to the world.

The increased opportunity to publish materials is certainly observable in the countless personal pages, as well as pages by families, small shops, etc., facilitated by the emergence of free Web hosting services.

Statistics

According to a 2001 study, there were more than 550 million documents on the Web, mostly in the "invisible Web". A 2002 survey of 2,024 million Web pages determined that by far the most Web content was in English: 56.4%; next were pages in German (7.7%), French (5.6%) and Japanese (4.9%). A more recent study which used web searches in 75 different languages to sample the Web determined that there were over 11.5 billion web pages in the publicly indexable Web as of January 2005.

Speed Issues

Frustration over congestion issues in the Internet infrastructure and the high latency that results in slow browsing has led to an alternative name for the World Wide Web: the World Wide Wait. Speeding up the Internet is an ongoing discussion over the use of peering and QoS technologies. Other solutions to reduce the World Wide Wait can be found on W3C.

Standard guidelines for ideal Web response times are (Nielsen 1999, page 42):

- 0.1 second (one tenth of a second). Ideal response time. The user doesn't sense any interruption.
- 1 second. Highest acceptable response time. Download times above 1 second interrupt the user experience.
- 10 seconds. Unacceptable response time. The user experience is interrupted and the user is likely to leave the site or system.

These numbers are useful for planning server capacity.

Link Rot

The Web suffers from link rot, links becoming broken because of the continual disappearance or relocation of Web resources over time. The ephemeral nature of the Web has prompted many efforts to archive the Web. The Internet Archive is one of the most well-known efforts; they have been archiving the Web since 1996.

Ultimately, link rot is the price paid for the World Wide Web; the web's scale and infrastructure make it impossible to monitor all hyperlinks in real time, and there is no broadcast notification when a page is removed or renamed. Allowing links between independently maintained resources carries the inherent danger that some of the links will be invalidated over time.

Academic Conferences

The major academic event covering the WWW is the World Wide Web series of conferences, promoted by IW3C2. There is a list with links to all conferences in the series.

WWW Prefix In Web Addresses

"WWW" is commonly found at the beginning of Web addresses because many organizations in the past followed a convention of naming hosts (servers) according to the services they provide. So for example, the host name for a Web server was often "www"; for an FTP server, "ftp"; and for a news server, "news" or "nntp" (after the news protocol NNTP). These host names then appeared as DNS subdomain names, as in "www.example.com".

This use of host or subdomain names and therefore such prefixes are not required by any technical standard; indeed, the first Web server was at "nxoc01.cern.ch" and even today many Web sites are available without a "www" prefix. The "www" prefix has no meaning in the way the main website is shown; locally or around the world. The "www" prefix is simply one choice for a Web site's subdomain name.

Some Web browsers will automatically try adding "www." to the beginning, and possibly ".com" to the end, of typed URLs if no host is found without them. Internet Explorer and Mozilla Firefox will also prefix "http://www." and append ".com" to the address bar contents if the Control and Enter keys are pressed simultaneously. For example, entering "example" in the address bar and then pressing either just Enter or Control+Enter will usually resolve to "http://www.example.com", depending on the exact browser version and its settings.

Pronunciation Of "WWW"

In English, WWW is the longest possible three-letter acronym (TLA) to pronounce, requiring nine syllables. The late Douglas Adams once quipped:

"The World Wide Web is the only thing I know of whose shortened form takes three times longer to say than its long form." (Douglas Adams, The Independent on Sunday, 1999)

Shorter variants include "triple double u", "three dub", "triple dub", "all the double u's", and the most common "dub-yu dub-yu dub-yu" or "dub dub dub". In other languages, "www" is often pronounced as "vvv" or "3w". The early "w^3" abbreviation is now defunct.

In Chinese, the World Wide Web is commonly translated to *wàn wéi wǎng*, i.e. "ten-thousand dimensional net".

Standards

The following is a cursory list of the documents that define the World Wide Web's three core standards:

Uniform Resource Locators (URL)

- RFC 1738, Uniform Resource Locators (URL) (December 1994)
- RFC 3986, Uniform Resource Identifier (URI): Generic Syntax (January 2005)

HyperText Transfer Protocol (HTTP)

- RFC 1945, HTTP/1.0 specification (May 1996)
- RFC 2616, HTTP/1.1 specification (June 1999)
- RFC 2617, HTTP Authentication
- HTTP/1.1 specification errata

Hypertext Markup Language (HTML)

- Internet Draft, HTML version 1
- RFC 1866, HTML version 2.0
- HTML 3.2 Reference Specification
- HTML 4.01 Specification
- Extensible HTML (XHTML) Specification

The Internet – URLs

What Is It?

Uniform Resource Locator (URL) is a technical, Web-related term used in two distinct meanings:

- in popular usage, it is a widespread synonym for Uniform Resource Identifier (URI) — many popular and technical texts will use the term "URL" when referring to URI;
- strictly, the idea of a uniform syntax for global identifiers of network-retrievable documents was the core idea of the World Wide Web. In the early times, these identifiers were variously called "document names", "Web addresses" and "Uniform Resource Locators". These names were misleading, however, because not all identifiers were locators, and even for those that were, this was not their defining characteristic. Nevertheless, by the time the RFC 1630 formally defined the term "URI" as a generic term best suited to the concept, the term "URL" had gained widespread popularity, which has continued to this day.

URI/URL Syntax In Brief

Here is a typical URI dissected:

```
http://user:pass@example.com:992/animal/bird?species=seagull#wings
\___/\_________/\____________/\__/\___________/\____________/\_____/
   |        |           |          |        |              |          |
protocol  login      hosts        port     path          query    anchor/fragment
```

Every URI (and therefore every URL) begins with the scheme name that defines its namespace, purpose, and the syntax of the remaining part of the URI. Most Web-enabled programs will try to dereference a URI according to the semantics of its scheme and a context-specific heuristic. For example, a Web browser will usually dereference a http://example.org/ by performing an HTTP request to the host example.org, at the default HTTP port (see Port 80). Dereferencing URI mailto:bob@example.com will usually open a "Compose e-mail" window with the address bob@example.com in the "To" field.

"example.com" is a domain name; an IP address or other network address might be used instead.

URLs As Locators

In its current strict technical meaning, a URL is a URI that, “in addition to identifying a resource, [provides] a means of locating the resource by describing its primary access mechanism (e.g., its network ‘location’).”

Clean URLs

"Clean" and "cruft-free" describe URLs which are:

- Not tied to technical details, such as the software used or whether the resource comes from a file or a database - so that a change in the technology will not break

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited.**io

existing links to the resource. e.g. /cars/audi/ is preferable to /cars/audi/index.php or /myprog.jsp?page=cars/audi/.

- Not tied to internal organisational structure, such as the current editor or department that created the document - so an internal reorganisation will not cause existing links to the document to break. e.g. /recommendations/2007/xyz/ is better than /~users/jane/current-work/xyz/ or /xyz-team/recommendations/.

- Consistent with other URLs in the same site in terms of hierarchy. This is desirable so a user can see where they are in the structure of the site, and can predict where to find what they are looking for. e.g. /cars/audi/ and /cars/ford/, instead of /cars/audi/ but /ford-cars/.

- Consistent with other URLs in the same site in terms of action. This is desirable so a user can predict other, similar URLs on that site, e.g. if /blogs/andrea/rss/ shows an RSS feed of Andrea's blog, then appending /rss/ to any another blog on the same site should show an RSS feed for that blog.

- A single location for a single resource. The same resource should not be available from multiple URLs, as this results in both confusion (Are they the same resource, or is one a copy of the other? Which is the 'right' one? Is one new and the other due to be removed?) and technical difficulties, e.g. counting links to a particular resource, or caching content to speed up access but not being able to show the cached content when the resource is accessed using a different URL.

An example of the difference between "clean" and "standard" URLs could be seen as:

Standard

http://example.com/index.php?section=articles&subsection=recent

Clean

http://example.com/articles/recent/

or

http://example.com/articles/2007/

Clean URLs With Web Services

Web services have been created that allow users to create short URLs which are easier to write down, remember or pass around. They are also more suitable for use where space is limited, for example in an IRC conversation, email signature, online forum or fixed width document (eg. email). A sample of current web services are provided below

- TinyURL.com - probably the most widely used due to its memorable name. Example URL http://www.tinyurl.com/2unsh
- doiop.com - one of the early services which offers keywords as opposed to random urls. Example doiop.com/keyword
- SnipURL.com
- shorl.com

Criticisms

Ultimately these services hide the ultimate destination from a web user. This can be used to unwittingly send people to sites that offend their sensibilities, or crash or compromise their computer using browser vulnerabilities. To help combat such abuse, TinyURL allows a user to set a cookie-based preference such that TinyURL stops at the TinyURL website, giving a preview of the final link, when that user clicks TinyURLs. Substituting preview.tinyurl.com for tinyurl.com in the URL is another way of stopping at a preview of the final link before clicking through to it. Opaqueness is also leveraged by spammers, who can use such links in spam (mostly blog spam), bypassing URL blacklists.

Dependency on a third-party service that may change, go away, or maintain privacy-compromising logs of user activity indefinitely.

Abuse

Early on, the creation of TinyURL IDs was predictable, and therefore could be exploited by users to create vulgar associations. For example, http://tinyurl.com/dick was made to link to the White House website of Dick Cheney. It now returns a message apologizing for any offense. Services that allow user's to define keywords (which are included in the URL) are more directly open to abuse since the user can essentially define any URL they chose.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www*.planetlimited.*io

The Internet – Domain Names

What Is It?

The term domain name has multiple related meanings:

- A name that identifies a computer or computers on the internet. These names appear as a component of a Web site's URL, e.g. www.wikipedia.org. This type of domain name is also called a hostname.
- The product that Domain name registrars provide to their customers. These names are often called registered domain names.
- Names used for other purposes in the Domain Name System (DNS), for example the special name which follows the @ sign in an email address, or the Top-level domains like .com, or the names used by the Session Initiation Protocol (VoIP), or DomainKeys.

They are sometimes colloquially (and incorrectly) referred to by marketers as "web addresses".

This article will primarily discuss registered domain names. See the Domain Name System article for technical discussions about general domain names and the hostname article for further information about the most common type of domain name.

Overview

The most common types of domain names are hostnames that provide more memorable names to stand in for numeric IP addresses. They allow for any service to move to a different location in the topology of the Internet (or an intranet), which would then have a different IP address.

By allowing the use of unique alphabetical addresses instead of numeric ones, domain names allow Internet users to more easily find and communicate with web sites and other server-based services. The flexibility of the domain name system allows multiple IP addresses to be assigned to a single domain name, or multiple domain names to be assigned to a single IP address. This means that one server may have multiple roles (such as hosting multiple independent Web sites), or that one role can be spread among many servers. One IP address can also be assigned to several servers, as used in anycast and hijacked IP space.

Hostnames are restricted to the ASCII letters "a" through "z' (case-insensitive), the digits "0" through "9", and the hyphen, with some other restrictions. Registrars restrict the domains to valid hostnames, since, otherwise, they would be useless. The Internationalized domain name (IDN) system has been developed to bypass the restrictions on character allowances in hostnames, making it easier for non-english alphabets to use the Internet. The underscore character is frequently used to ensure that a domain name is not recognized as a hostname, for example with the use of SRV records, although some older systems, such as NetBIOS did allow it. Due to confusion and other reasons, domain names with underscores in them are sometimes used where hostnames are required.

Examples

The following example illustrates the difference between a URL (Uniform Resource Locator) and a domain name:

- URL: http://www.example.net/index.html
- Domain name: www.example.net
- Registered domain name: example.net

As a general rule, the IP address and the server name are interchangeable. For most Internet services, the server will not have any way to know which was used. However, the explosion of interest in the Web means that there are far more Web sites than servers. To accommodate this, the hypertext transfer protocol (HTTP) specifies that the client tells the server which name is being used. This way, one server with one IP address can provide different sites for different domain names. This feature goes under the name virtual hosting and is commonly used by Web hosts.

For example, as referenced in RFC 2606 (Reserved Top Level DNS Names), the server at IP address 192.0.34.166 handles all of the following sites:

- example.com
- www.example.com
- example.net
- www.example.net
- example.org
- www.example.org

When a request is made, the data corresponding to the hostname requested is served to the user.

Top-Level Domains

Every domain name ends in a top-level domain (TLD) name, which is always either one of a small list of generic names (three or more characters), or a two characters territory code based on ISO-3166 (there are few exceptions and new codes are integrated case by case). Top-level domains are sometimes also called first-level domains.

The generic top-level domain (gTLD) extensions are:

- .aero - for the air transport industry
- .biz - for business use
- .cat - for Catalan language/culture
- .com - for commercial organizations, but unrestricted
- .coop - for cooperatives
- .edu - for post-secondary educational establishments
- .gov - for governments and their agencies in the United States
- .info - for informational sites, but unrestricted
- .int - for international organizations established by treaty
- .jobs - for employment-related sites
- .mil - for the US military
- .mobi - for sites catering to mobile devices
- .museum - for museums
- .name - for families and individuals
- .net - originally for network infrastructures, now unrestricted

- .org - originally for organizations not clearly falling within the other gTLDs, now unrestricted
- .pro - for certain professions
- .tel - for services involving connections between the telephone network and the Internet (added March 2, 2007)
- .travel - for travel agents, airlines, hoteliers, tourism bureaus, etc.

The following gTLDs are in the process of being approved, and may be added to the root nameservers in the near future:

- .asia - for the Asian community
- .post - for postal services
- .geo - for geographically related sites

The country code top-level domain (ccTLD) extensions are:

(Please note: * = Foreign registration permitted)

- .ac – Ascension Island *
- .ad – Andorra
- .ae – United Arab Emirates
- .af – Afghanistan
- .ag – Antigua and Barbuda *
- .ai – Anguilla
- .al – Albania
- .am – Armenia *
- .an – Netherlands Antilles
- .ao – Angola
- .aq – Antarctica
- .ar – Argentina
- .as – American Samoa *
- .at – Austria *
- .au – Australia
- .aw – Aruba
- .ax – Aland Islands
- .az – Azerbaijan
- .ba – Bosnia and Herzegovina
- .bb – Barbados
- .bd – Bangladesh
- .be – Belgium *
- .bf – Burkina Faso
- .bg – Bulgaria
- .bh – Bahrain
- .bi – Burundi *
- .bj – Benin
- .bm – Bermuda
- .bn – Brunei
- .bo – Bolivia *
- .br – Brazil *
- .bs – Bahamas *
- .bt – Bhutan
- .bu – Burma (not in use since re-naming of country to Myanmar, see .mm)
- .bv – Bouvet Island (not in use; no registrations)

- .bw – Botswana
- .by – Belarus
- .bz – Belize *
- .ca – Canada
- .cc – Cocos (Keeling) Islands *
- .cd – Democratic Republic of the Congo (formerly .zr – Zaire) *
- .cf – Central African Republic
- .cg – Republic of the Congo *
- .ch – Switzerland *
- .ci – Côte d'Ivoire (Ivory Coast)
- .ck – Cook Islands *
- .cl – Chile
- .cm – Cameroon
- .cn – People's Republic of China *
- .co – Colombia
- .cr – Costa Rica
- .cs – Serbia and Montenegro (formerly .yu – Yugoslavia; Note: on June 3, 2006, Montenegro declared independence, thus dissolving the state union) (.cs code not assigned; no DNS) (.cs code previously used for Czechoslovakia)
- .cu – Cuba
- .cv – Cape Verde
- .cx – Christmas Island *
- .cy – Cyprus
- .cz – Czech Republic
- .de – Germany
- .dj – Djibouti *
- .dk – Denmark *
- .dm – Dominica
- .do – Dominican Republic
- .dz – Algeria
- .ec – Ecuador *
- .ee – Estonia
- .eg – Egypt
- .eh – Western Sahara (not assigned; no DNS)
- .er – Eritrea
- .es – Spain *
- .et – Ethiopia
- .eu – European Union (code "exceptionally reserved" by ISO 3166-1)
- .fi – Finland
- .fj – Fiji *
- .fk – Falkland Islands
- .fm – Federated States of Micronesia *
- .fo – Faroe Islands
- .fr – France
- .ga – Gabon
- .gb – United Kingdom (Reserved domain by IANA; deprecated – see .uk)
- .gd – Grenada
- .ge – Georgia
- .gf – French Guiana
- .gg – Guernsey
- .gh – Ghana
- .gi – Gibraltar
- .gl – Greenland *

- .gm – Gambia
- .gn – Guinea
- .gp – Guadeloupe
- .gq – Equatorial Guinea
- .gr – Greece *
- .gs – South Georgia and the South Sandwich Islands *
- .gt – Guatemala
- .gu – Guam
- .gw – Guinea-Bissau
- .gy – Guyana
- .hk – Hong Kong *
- .hm – Heard Island and McDonald Islands *
- .hn – Honduras *
- .hr – Croatia
- .ht – Haiti
- .hu – Hungary *
- .id – Indonesia
- .ie – Ireland
- .il – Israel *
- .im – Isle of Man *
- .in – India *
- .io – British Indian Ocean Territory *
- .iq – Iraq
- .ir – Iran *
- .is – Iceland
- .it – Italy
- .je – Jersey
- .jm – Jamaica
- .jo – Jordan
- .jp – Japan
- .ke – Kenya
- .kg – Kyrgyzstan
- .kh – Cambodia
- .ki – Kiribati
- .km – Comoros
- .kn – Saint Kitts and Nevis
- .kp – North Korea (not assigned; no DNS)
- .kr – South Korea
- .kw – Kuwait
- .ky – Cayman Islands
- .kz – Kazakhstan *
- .la – Laos *
- .lb – Lebanon
- .lc – Saint Lucia
- .li – Liechtenstein *
- .lk – Sri Lanka
- .lr – Liberia
- .ls – Lesotho
- .lt – Lithuania
- .lu – Luxembourg
- .lv – Latvia *
- .ly – Libya *
- .ma – Morocco

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited*****.io***

- .mc – Monaco
- .md – Moldova *
- .me – Montenegro
- .mg – Madagascar
- .mh – Marshall Islands
- .mk – Republic of Macedonia
- .ml – Mali
- .mm – Myanmar (formerly .bu – Burma)
- .mn – Mongolia *
- .mo – Macau
- .mp – Northern Mariana Islands *
- .mq – Martinique
- .mr – Mauritania
- .ms – Montserrat *
- .mt – Malta
- .mu – Mauritius *
- .mv – Maldives
- .mw – Malawi *
- .mx – Mexico *
- .my – Malaysia
- .mz – Mozambique
- .na – Namibia *
- .nc – New Caledonia
- .ne – Niger
- .nf – Norfolk Island *
- .ng – Nigeria
- .ni – Nicaragua
- .nl – Netherlands * (first ccTLD registered)
- .no – Norway
- .np – Nepal
- .nr – Nauru *
- .nu – Niue *
- .nz – New Zealand *
- .om – Oman
- .pa – Panama
- .pe – Peru
- .pf – French Polynesia
- .pg – Papua New Guinea
- .ph – Philippines *
- .pk – Pakistan *
- .pl – Poland *
- .pm – Saint Pierre and Miquelon
- .pn – Pitcairn Islands *
- .pr – Puerto Rico *
- .ps – Palestinian territories *
- .pt – Portugal *
- .pw – Palau
- .py – Paraguay
- .qa – Qatar
- .re – Réunion
- .ro – Romania *
- .rs – Serbia
- .ru – Russia *

- .rw – Rwanda
- .sa – Saudi Arabia
- .sb – Solomon Islands *
- .sc – Seychelles *
- .sd – Sudan
- .se – Sweden *
- .sg – Singapore
- .sh – Saint Helena *
- .si – Slovenia
- .sj – Svalbard and Jan Mayen islands (not in use; no registrations)
- .sk – Slovakia
- .sl – Sierra Leone
- .sm – San Marino *
- .sn – Senegal
- .so – Somalia
- .sr – Suriname *
- .st – São Tomé and Príncipe *
- .su – Soviet Union (deprecated; being phased out; code "transitionally reserved" by ISO 3166-1)
- .sv – El Salvador
- .sy – Syria *
- .sz – Swaziland *
- .tc – Turks and Caicos Islands
- .td – Chad
- .tf – French Southern Territories
- .tg – Togo *
- .th – Thailand
- .tj – Tajikistan *
- .tk – Tokelau *
- .tl – East Timor (formerly .tp) *
- .tm – Turkmenistan *
- .tn – Tunisia
- .to – Tonga *
- .tp – East Timor (deprecated – use .tl; code "transitionally reserved" by ISO 3166-1)
- .tr – Turkey
- .tt – Trinidad and Tobago *
- .tv – Tuvalu *
- .tw – Republic of China (Taiwan) *
- .tz – Tanzania
- .ua – Ukraine
- .ug – Uganda *
- .uk – United Kingdom (code "exceptionally reserved" by ISO 3166-1) (see also .gb)
- .us – United States *
- .uy – Uruguay
- .uz – Uzbekistan
- .va – Vatican City
- .vc – Saint Vincent and the Grenadines *
- .ve – Venezuela
- .vg – British Virgin Islands *
- .vi – United States Virgin Islands
- .vn – Vietnam
- .vu – Vanuatu *
- .wf – Wallis and Futuna

- .ws – Samoa (formerly Western Samoa) *
- .ye – Yemen
- .yt – Mayotte
- .yu – Yugoslavia (subsequently renamed Serbia and Montenegro - code officially replaced by .cs (see above) but still used; code "transitionally reserved" by ISO 3166-1)
- .za – South Africa *
- .zm – Zambia
- .zw – Zimbabwe

Other-Level Domains

In addition to the top-level domains, there are second-level domain (SLD) names. These are the names directly to the left of .com, .net, and the other top-level domains. As an example, in the domain en.wikipedia.org, "wikipedia" is the second-level domain.

On the next level are third-level domains. These domains are immediately to the left of a second-level domain. In the en.wikipedia.org example, "en" is a third-level domain. There can be fourth and fifth level domains and so on, with virtually no limitation. An example of a working domain with five levels is www.sos.state.oh.us. Each level is separated by a dot or period symbol between them.

Domains of third or higher level are also known as subdomains, though this term technically applies to a domain of any level, since even a top-level domain is a "subdomain" of the "root" domain (a "zeroth-level" domain that is designated by a dot alone).

Traditionally, the second level domain was the name of the company or the name used on the internet. The third level was commonly used to designate a particular host server. Therefore, ftp.wikipedia.org might be an FTP server, www.wikipedia.org would be a World Wide Web Server, and mail.wikipedia.org could be an email server. Modern technology now allows multiple servers to serve a single subdomain, or multiple protocols or domains to be served by a single computer. Therefore, subdomains may or may not have any real purpose.

Official Assignment

ICANN (Internet Corporation for Assigned Names and Numbers) has overall responsibility for managing the DNS. It controls the root domain, delegating control over each top-level domain to a domain name registry. For ccTLDs, the domain registry is typically controlled by the government of that country. ICANN has a consultation role in these domain registries but is in no position to regulate the terms and conditions of how a domain name is allocated or who allocates it in each of these country level domain registries. On the other hand, generic top-level domains (gTLDs) are governed directly under ICANN which means all terms and conditions are defined by ICANN with the cooperation of the gTLD registries.

Domain names which are theoretically leased can be considered in the same way as real estate, due to a significant impact on online brand building, advertising, search engine optimization, etc.

A few companies have offered low-cost, below-cost or even free domain registrations, with a variety of models adopted to recoup the costs to the provider. These usually require that domains are hosted on their site in a framework or portal, with advertising wrapped around the user's content, revenue from which allows the provider to recoup the costs. When the DNS was new, domain registrations were free. A domain owner can generally give away or sell infinite subdomains of their domain, e.g. the owner of example.edu could provide domains that are subdomains, such as foo.example.edu and foo.bar.example.edu.

Uses And Abuses

As domain names became attractive to marketers, rather than just the technical audience for which they were originally intended, they began to be used in manners that in many cases did not fit in their intended structure. As originally planned, the structure of domain names followed a strict hierarchy in which the top level domain indicated the type of organization (commercial, governmental, etc.), and addresses would be nested down to third, fourth, or further levels to express complex structures, where, for instance, branches, departments, and subsidiaries of a parent organization would have addresses which were subdomains of the parent domain. Also, hostnames were intended to correspond to actual physical machines on the network, generally with only one name per machine.

However, once the World Wide Web became popular, site operators frequently wished to have memorable addresses, regardless of whether they fit properly in the structure; thus, since the .com domain was the most popular and memorable, even noncommercial sites would often get addresses under it, and sites of all sorts wished to have second-level domain registrations even if they were parts of a larger entity where a logical subdomain would have made sense (e.g., abcnews.com instead of news.abc.com). A Web site found at http://www.example.org/ will often be advertised without the "http://", and in most cases can be reached by just entering "example.org" into a Web browser. In the case of a .com, the Web site can sometimes be reached by just entering "example" (depending on browser versions and configuration settings, which vary in how they interpret incomplete addresses).

The popularity of domain names also led to uses which were regarded as abusive by established companies with trademark rights; this was known as cybersquatting, in which somebody took a name that resembled a trademark in order to profit from traffic to that address. To combat this, various laws and policies were enacted to allow abusive registrations to be forcibly transferred, but these were sometimes themselves abused by overzealous companies committing reverse domain hijacking against domain users who had legitimate grounds to hold their names, such as their being generic words as well as trademarks in a particular context, or their use in the context of fan or protest sites with free speech rights of their own.

Laws that specifically address domain name conflicts include the Anticybersquatting Consumer Protection Act in the United States and the Trademarks Act, 1999, in India. Alternatively, domain registrants are bound by contract under the UDRP to comply with mandatory arbitration proceedings should someone challenge their ownership of the domain name.

Generic Domain Names — Problems Arising Out Of Unregulated Name Selection

Within a particular top-level domain, parties are generally free to select an unallocated domain name as their own on a first come, first served basis, resulting in Harris's lament, all the good ones are taken. For generic or commonly used names, this may sometimes lead to the use of a domain name which is inaccurate or misleading. This problem can be seen with regard to the ownership or control of domain names for a generic product or service.

By way of illustration, there has been tremendous growth in the number and size of literary festivals around the world in recent years. In this context, currently a generic domain name such as literary.org is available to the first literary festival organisation which is able to obtain registration, even if the festival in question is very young or obscure. Some critics would argue that there is greater amenity in reserving such domain names for the use of, for example, a regional or umbrella grouping of festivals. Related issues may also arise in relation to non-commercial domain names.

Unconventional Domain Names

Due to the rarity of one-word dot-com domain names, many unconventional domain names, domain hacks, have been gaining popularity. They make use of the top-level domain as an integral part of the Web site's title. Two popular domain hack Web sites are del.icio.us and blo.gs, which spell out "delicious" and "blogs", respectively.

Unconventional domain names are also used to create unconventional email addresses. Non-working examples that spell 'James' are j@m.es and j@mes.com, which use the domain names m.es (of Spain's .es) and mes.com, respectively.

Commercial Resale Of Domain Names

An economic effect of the widespread usage of domain names has been the resale market for generic domain names that has sprung up in the last decade. Certain domains, especially those related to business, gambling, pornography, and other commercially lucrative fields of digital world trade have become very much in demand to corporations and entrepreneurs due to their intrinsic value in attracting clients. The most expensive Internet domain name to date, according to Guinness World Records, is business.com which was resold in 1999 for $7.5 million, but this was $7.5 million in stock options, not in cash. Later the stock was valued at, not sold, for $2 million and may even be worth less today. There are disputes about the high values of domain names claimed and the actual prices of many sales.

Another high value domain name, sex.com, was stolen from its rightful owner by means of a forged transfer instruction via fax. During the height of the dot-com era, the domain was earning millions of dollars per month in advertising revenue from the large influx of visitors that arrived daily. Two long-running U.S. lawsuits resulted, one against the thief and one against the domain registrar VeriSign. In one of the cases, Kremen v. Network Solutions, the court found in favor of the plaintiff, leading to an unprecedented ruling that classified domain names as property, granting them the same legal protections. In 1999, Microsoft traded the valuable name Bob.com with internet entrepreneur Bob Kerstein for the name Windows2000.com which was the name of their new operating system.

One of the reasons for the value of domain names is that even without advertising or marketing, they attract clients seeking services and products who simply type in the generic name. Furthermore, generic domain names such as movies.com or Books.com are extremely easy for potential customers to remember, increasing the probability that they become repeat customers or regular clients.

Although the current domain market is nowhere as strong as it was during the dot-com heyday, it remains strong and is currently experiencing solid growth again. Annually tens of millions of dollars change hands due to the resale of domains. Large numbers of registered domain names lapse and are deleted each year. On average 25,000 domain names drop (are deleted) every day.

People who buy and sell domain names are known as domainers.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

Domain Name Confusion

Intercapping is often used to clarify a domain name. However, DNS is case-insensitive, and some names may be misinterpreted when converted to lowercase. For example: Who Represents, a database of artists and agents, chose whorepresents.com; a therapists' network thought therapistfinder.com looked good; and another website operating as of October, 2006, is penisland.net a website for Pen Island, a site that claims to be an online pen vendor, but exists primarily as a joke, as it has no products for sale. In such situations, the proper wording can be clarified by use of hyphens. For instance, Experts Exchange, the programmers' site, for a long time used expertsexchange.com, but ultimately changed the name to experts-exchange.com.

Leo Stoller threatened to sue the owners of StealThisEmail.com on the basis that, when read as StealthIsEmail.com, it infringed on claimed trademark rights to the word "stealth".

The Internet – The Domain Name System

What Is It?

On the Internet, the Domain Name Service (DNS) stores and associates many types of information with domain names; most importantly, it translates domain names (computer hostnames) to IP addresses. It also lists mail exchange servers accepting e-mail for each domain. In providing a worldwide keyword-based redirection service, DNS is an essential component of contemporary Internet use.

Pre-eminently, the DNS makes it possible to assign Internet destinations to the human organization or concern they represent, indepenently of the physical routing hierarchy represented by the numerical IP address. Because of this, hyperlinks and Internet contact information can remain the same, whatever the current IP routing arrangements may be, and can take a human-readable form (such as "wikipedia.org") which is rather easier to remember than an IP address (such as 66.230.200.100). People take advantage of this when they recite meaningful URLs and e-mail addresses without caring how the machine will actually locate them.

The DNS also distributes the responsibility for assigning domain names and mapping them to IP networks by allowing an authoritative server for each domain to keep track of its own changes, avoiding the need for a central registrar to be continually consulted and updated.

The Five Layer TCP/IP Model

5. Application Layer

DHCP • DNS • FTP • HTTP • IMAP4 • IRC • NNTP • XMPP • MIME • POP3 • SIP • SMTP • SNMP • SSH • TELNET • BGP • RPC • RTP • RTCP • TLS/SSL • SDP • SOAP • L2TP • PPTP • …

4. Transport Layer

TCP • UDP • DCCP • SCTP • GTP • …

3. Network Layer

IP (IPv4 • IPv6) • ICMP • IGMP • RSVP • IPsec • …

2. Data link Layer

ATM • DTM • Ethernet • FDDI • Frame Relay • GPRS • PPP • ARP • RARP • …

1. Physical Layer

Ethernet physical layer • ISDN • Modems • PLC • SONET/SDH • G.709 • Wi-Fi • …

History Of The DNS

The practice of using a name as a more human-legible abstraction of a machine's numerical address on the network predates even TCP/IP, and goes all the way to the ARPAnet era.

Originally, each computer on the network retrieved a file called HOSTS.TXT from SRI (now SRI International) which mapped an address (such as 192.0.34.166) to a name (such as www.example.net.) The Hosts file still exists on most modern operating systems, either by default or through configuration, and allows users to specify an IP address to use for a hostname without checking the DNS. This file now serves primarily for troubleshooting DNS errors or for mapping local addresses to more organic names. Systems based on a HOSTS.TXT file have inherent limitations, because of the obvious requirement that every time a given computer's address changed, every computer that seeks to communicate with it would need an update to its Hosts file.

The growth of networking called for a more scalable system: one that recorded a change in a host's address in one place only. Other hosts would learn about the change dynamically through a notification system, thus completing a globally accessible network of all hosts' names and their associated IP Addresses.

Paul Mockapetris invented the DNS in 1983 and wrote the first implementation. The original specifications appear in RFC 882 and 883. In 1987, the publication of RFC 1034 and RFC 1035 updated the DNS specification and made RFC 882 and RFC 883 obsolete. Several more-recent RFCs have proposed various extensions to the core DNS protocols.

In 1984, four Berkeley students — Douglas Terry, Mark Painter, David Riggle and Songnian Zhou — wrote the first UNIX implementation, which was maintained by Ralph Campbell thereafter. In 1985, Kevin Dunlap of DEC significantly re-wrote the DNS implementation and renamed it BIND (Berkeley Internet Name Domain, previously: Berkeley Internet Name Daemon). Mike Karels, Phil Almquist and Paul Vixie have maintained BIND since then. BIND was ported to the Windows NT platform in the early 1990s.

Due to its long history of security issues, several alternative nameserver/resolver programs have been written and distributed in recent years.

How The DNS Works In Theory

The domain name space consists of a tree of domain names. Each node or leaf in the tree has one or more resource records, which hold information associated with the domain name. The tree sub-divides into zones. A zone consists of a collection of connected nodes authoritatively served by an authoritative DNS nameserver. (Note that a single nameserver can host several zones.)

When a system administrator wants to let another administrator control a part of the domain name space within his or her zone of authority, he or she can delegate control to the other administrator. This splits a part of the old zone off into a new zone, which comes under the authority of the second administrator's nameservers. The old zone becomes no longer authoritative for what comes under the authority of the new zone.

A resolver looks up the information associated with nodes. A resolver knows how to communicate with name servers by sending DNS requests, and heeding DNS responses. Resolving usually entails iterating through several name servers to find the needed information.

Some resolvers function simplistically and can only communicate with a single name server. These simple resolvers rely on a recursing name server to perform the work of finding information for them.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited.**io

Understanding The Parts Of A Domain Name

A domain name usually consists of two or more parts (technically labels), separated by dots. For example wikipedia.org.

- The rightmost label conveys the top-level domain (for example, the address en.wikipedia.org has the top-level domain org).
- Each label to the left specifies a subdivision or subdomain of the domain above it. Note that "subdomain" expresses relative dependence, not absolute dependence: for example, wikipedia.org comprises a subdomain of the org domain, and en.wikipedia.org comprises a subdomain of the domain wikipedia.org. In theory, this subdivision can go down to 127 levels deep, and each label can contain up to 63 characters, as long as the whole domain name does not exceed a total length of 255 characters. But in practice some domain registries have shorter limits than that.
- A hostname refers to a domain name that has one or more associated IP addresses. For example, the en.wikipedia.org and wikipedia.org domains are both hostnames, but the org domain is not.

The DNS consists of a hierarchical set of DNS servers. Each domain or subdomain has one or more authoritative DNS servers that publish information about that domain and the name servers of any domains "beneath" it. The hierarchy of authoritative DNS servers matches the hierarchy of domains. At the top of the hierarchy stand the root servers: the servers to query when looking up (resolving) a top-level domain name (TLD).

The Address Resolution Mechanism

(Please note: This description deliberately uses the fictional .example TLD in accordance with the DNS guidelines themselves.)

In theory a full host name may have several name segments, (e.g ahost.ofasubnet.ofabiggernet.inadomain.example). In practice, in the experience of the majority of public users of Internet services, full host names will frequently consist of just three segments (ahost.inadomain.example, and most often www.inadomain.example).

For querying purposes, software interprets the name segment by segment, from right to left, using an iterative search procedure. At each step along the way, the program queries a corresponding DNS server to provide a pointer to the next server which it should consult.

As originally envisaged, the process was as simple as:

1. the local system is pre-configured with the known addresses of the root servers in a file of root hints, which need to be updated periodically by the local administrator from a reliable source to be kept up to date with the changes which occur over time.
2. query one of the root servers to find the server authoritative for the next level down (so in the case of our simple hostname, a root server would be asked for the address of a server with detailed knowledge of the example top level domain).
3. querying this second server for the address of a DNS server with detailed knowledge of the second-level domain (inadomain.example in our example).
4. repeating the previous step to progress down the name, until the final step which would, rather than generating the address of the next DNS server, return the final address sought.

The mechanism in this simple form has a difficulty: it places a huge operating burden on the collective of root servers, with each and every search for an address starting by querying one of them. Being as critical as they are to the overall function of the system such heavy use would create an insurmountable bottleneck for trillions of queries placed every day. In practice there are two key additions to the mechanism.

- Firstly, the DNS resolution process allows for local recording and subsequent consultation of the results of a query (or caching) for a period of time after a successful answer (the server providing the answer initially dictates the period of validity, which may vary from just seconds to days or even weeks). In our illustration, having found a list of addresses of servers capable of answering queries about the .example domain, the local resolver will not need to make the query again until the validity of the currently known list expires, and so on for all subsequent steps. Hence having successfully resolved the address of ahost.inadomain.example it is not necessary to repeat the process for some time since the address already reached will be deemed reliable for a defined period, and resolution of anotherhost.anotherdomain.example can commence with already knowing which servers can answer queries for the .example domain. Caching significantly reduces the rate at which the most critical name servers have to respond to queries, adding the extra benefit that subsequent resolutions are not delayed by network transit times for the queries and responses.

- Secondly, most domestic and small-business clients "hand off" address resolution to their ISP's DNS servers to perform the look-up process, thus allowing for the greatest benefit from those same ISPs having busy local caches serving a wide variety of queries and a large number of users.

Circular Dependencies And Glue Records

Name servers in delegations appear listed by name, rather than by IP address. This means that a resolving name server must issue another DNS request to find out the IP address of the server to which it has been referred. Since this can introduce a circular dependency if the nameserver referred to is under the domain that it is authoritative of, it is occasionally necessary for the nameserver providing the delegation to also provide the IP address of the next nameserver. This record is called a glue record.

For example, assume that the sub-domain en.wikipedia.org contains further sub-domains (such as something.en.wikipedia.org) and that the authoritative nameserver for these lives at ns1.en.wikipedia.org. A computer trying to resolve something.en.wikipedia.org will thus first have to resolve ns1.en.wikipedia.org. Since ns1 is also under the en.wikipedia.org subdomain, resolving ns1.en.wikipedia.org requires resolving ns1.en.wikipedia.org which is exactly the circular dependency mentioned above. The dependency is broken by the glue record in the nameserver of wikipedia.org that provides the IP address of ns1.en.wikipedia.org directly to the requestor, enabling it to bootstrap the process by figuring out where ns1.en.wikipedia.org is located.

DNS In Practice

When an application (such as a web browser) tries to find the IP address of a domain name, it doesn't necessarily follow all of the steps outlined in the Theory section above. We will first look at the concept of caching, and then outline the operation of DNS in "the real world."

Caching And Time To Live

Because of the huge volume of requests generated by a system like the DNS, the designers wished to provide a mechanism to reduce the load on individual DNS servers. The mechanism devised provided that when a DNS resolver (i.e. client) received a DNS response, it would cache that response for a given period of time. A value (set by the administrator of the DNS server handing out the response) called the time to live (TTL), defines that period of time. Once a response goes into cache, the resolver will consult its cached (stored) answer; only when the TTL expires (or when an administrator manually flushes the response from the resolver's memory) will the resolver contact the DNS server for the same information.

Generally, the Start of Authority (SOA) record specifies the time to live. The SOA record has the parameters:

- Serial — the zone serial number, incremented when the zone file is modified, so the slave and secondary name servers know when the zone has been changed and should be reloaded.
- Refresh — the number of seconds between update requests from secondary and slave name servers.
- Retry — the number of seconds the secondary or slave will wait before retrying when the last attempt has failed.
- Expire — the number of seconds a master or slave will wait before considering the data stale if it cannot reach the primary name server.
- Minimum — previously used to determine the minimum TTL, this offers negative caching.

(Newer versions of BIND (named) will accept the suffixes 'M','H','D' or 'W', indicating a time-interval of minutes, hours, days and weeks respectively.)

Caching Time

As a noteworthy consequence of this distributed and caching architecture, changes to the DNS do not always take effect immediately and globally. This is best explained with an example: If an administrator has set a TTL of 6 hours for the host www.wikipedia.org, and then changes the IP address to which www.wikipedia.org resolves at 12:01pm, the administrator must consider that a person who cached a response with the old IP address at 12:00pm will not consult the DNS server again until 6:00pm. The period between 12:01pm and 6:00pm in this example is called caching time, which is best defined as a period of time that begins when you make a change to a DNS record and ends after the maximum amount of time specified by the TTL expires. This essentially leads to an important logistical consideration when making changes to the DNS: not everyone is necessarily seeing the same thing you're seeing. RFC 1537 helps to convey basic rules for how to set the TTL.

Note that the term "propagation", although very widely used, does not describe the effects of caching well. Specifically, it implies that [1] when you make a DNS change, it somehow spreads to all other DNS servers (instead, other DNS servers check in with yours as needed), and [2] that you do not have control over the amount of time the record is cached (you control the TTL values for all DNS records in your domain, except your NS records and any authoritative DNS servers that use your domain name).

Some resolvers may override TTL values, as the protocol supports caching for up to 68 years or no caching at all. Negative caching (the non-existence of records) is determined by name servers authoritative for a zone which MUST include the SOA record when reporting no data of the requested type exists. The MINIMUM field of the SOA record and the TTL of the SOA itself is used to establish the TTL for the negative answer. RFC 2308

Many people incorrectly refer to a mysterious 48 hour or 72 hour propagation time when you make a DNS change. When one changes the NS records for one's domain or the IP addresses for hostnames of authoritative DNS servers using one's domain (if any), there can be a lengthy period of time before all DNS servers use the new information. This is because those records are handled by the zone parent DNS servers (for example, the .com DNS servers if your domain is example.com), which typically cache those records for 48 hours. However, those DNS changes will be immediately available for any DNS servers that do not have them cached. And, any DNS changes on your domain other than the NS records and authoritative DNS server names can be nearly instantaneous, if you choose for them to be (by lowering the TTL once or twice ahead of time, and waiting until the old TTL expires before making the change).

DNS In The Real World

Users generally do not communicate directly with a DNS resolver. Instead DNS resolution takes place transparently in client applications such as web browsers (like Opera, Mozilla Firefox, Safari, Netscape Navigator, Internet Explorer, etc.), mail clients (Outlook Express, Mozilla Thunderbird, etc.), and other Internet applications. When a request is made which necessitates a DNS lookup, such programs send a resolution request to the local DNS resolver in the operating system which in turn handles the communications required.

The DNS resolver will almost invariably have a cache (see above) containing recent lookups. If the cache can provide the answer to the request, the resolver will return the value in the cache to the program that made the request. If the cache does not contain the answer, the resolver will send the request to a designated DNS server or servers. In the case of most home users, the Internet service provider to which the machine connects will usually supply this DNS server: such a user will either configure that server's address manually or allow DHCP to set it; however, where systems administrators have configured systems to use their own DNS servers, their DNS resolvers will generally point to their own nameservers. This name server will then follow the process outlined above in DNS in theory, until it either successfully finds a result, or does not. It then returns its results to the DNS resolver; assuming it has found a result, the resolver duly caches that result for future use, and hands the result back to the software which initiated the request.

Broken Resolvers

An additional level of complexity emerges when resolvers violate the rules of the DNS protocol. Some people have suggested that a number of large ISPs have configured their DNS servers to violate rules (presumably to allow them to run on less-expensive hardware than a fully compliant resolver), such as by disobeying TTLs, or by indicating that a domain name does not exist just because one of its name servers does not respond.

As a final level of complexity, some applications such as Web browsers also have their own DNS cache, in order to reduce the use of the DNS resolver library itself. This practice can add extra difficulty to DNS debugging, as it obscures which data is fresh, or lies in which cache. These caches typically have very short caching times of the order of one minute. A notable exception is Internet Explorer; recent versions cache DNS records for half an hour

Other DNS Applications

The system outlined above provides a somewhat simplified scenario. The DNS includes several other functions:

- Hostnames and IP addresses do not necessarily match on a one-to-one basis. Many hostnames may correspond to a single IP address: combined with virtual hosting, this allows a single machine to serve many web sites. Alternatively a single hostname may correspond to many IP addresses: this can facilitate fault tolerance and load distribution, and also allows a site to move physical location seamlessly.
- There are many uses of DNS besides translating names to IP addresses. For instance, Mail transfer agents use DNS to find out where to deliver e-mail for a particular address. The domain to mail exchanger mapping provided by MX records accommodates another layer of fault tolerance and load distribution on top of the name to IP address mapping.
- Sender Policy Framework and DomainKeys instead of creating own record types were designed to take advantage of another DNS record type, the TXT record.
- To provide resilience in the event of computer failure, multiple DNS servers provide coverage of each domain. In particular, thirteen root servers exist worldwide. DNS programs or operating systems have the IP addresses of these servers built in. At least nominally, the USA hosts all but three of the root servers. However, because many root servers actually implement anycast, where many different computers can share the same IP address to deliver a single service over a large geographic region, most of the physical (rather than nominal) root servers now operate outside the USA.

The DNS uses TCP and UDP on port 53 to serve requests. Almost all DNS queries consist of a single UDP request from the client followed by a single UDP reply from the server. TCP typically comes into play only when the response data size exceeds 512 bytes, or for such tasks as zone transfer. Some operating systems such as HP-UX are known to have resolver implementations that use TCP for all queries, even when UDP would suffice.

Extensions To DNS

EDNS is an extension of the DNS protocol which enhances the transport of DNS data in UDP packages, and adds support for expanding the space of request and response codes. It is described in RFC 2671.

Implementations Of DNS

For a commented list of DNS server-side implementations, see Comparison of DNS server software.

Standards

- RFC 882 Concepts and Facilities (Deprecated by RFC 1034)
- RFC 883 Domain Names: Implementation specification (Deprecated by RFC 1035)
- RFC 1032 Domain administrators guide
- RFC 1033 Domain administrators operations guide
- RFC 1034 Domain Names - Concepts and Facilities.
- RFC 1035 Domain Names - Implementation and Specification
- RFC 1101 DNS Encodings of Network Names and Other Types
- RFC 1123 Requirements for Internet Hosts -- Application and Support
- RFC 1183 New DNS RR Definitions
- RFC 1706 DNS NSAP Resource Records
- RFC 1876 Location Information in the DNS (LOC)

- RFC 1886 DNS Extensions to support IP version 6
- RFC 1912 Common DNS Operational and Configuration Errors
- RFC 1995 Incremental Zone Transfer in DNS
- RFC 1996 A Mechanism for Prompt Notification of Zone Changes (DNS NOTIFY)
- RFC 2136 Dynamic Updates in the domain name system (DNS UPDATE)
- RFC 2181 Clarifications to the DNS Specification
- RFC 2182 Selection and Operation of Secondary DNS Servers
- RFC 2308 Negative Caching of DNS Queries (DNS NCACHE)
- RFC 2317 Classless IN-ADDR.ARPA delegation
- RFC 2671 Extension Mechanisms for DNS (EDNS0)
- RFC 2672 Non-Terminal DNS Name Redirection (DNAME record)
- RFC 2782 A DNS RR for specifying the location of services (DNS SRV)
- RFC 2845 Secret Key Transaction Authentication for DNS (TSIG)
- RFC 2874 DNS Extensions to Support IPv6 Address Aggregation and Renumbering
- RFC 3403 Dynamic Delegation Discovery System (DDDS) (NAPTR records)
- RFC 3696 Application Techniques for Checking and Transformation of Names
- RFC 4398 Storing Certificates in the Domain Name System
- RFC 4408 Sender Policy Framework (SPF) (SPF records)

Types Of DNS Records

Important categories of data stored in the DNS include the following:

- An A record or address record maps a hostname to a 32-bit IPv4 address.
- An AAAA record or IPv6 address record maps a hostname to a 128-bit IPv6 address.
- A CNAME record or canonical name record is an alias of one name to another. The A record that the alias is pointing to can be either local or remote - on a foreign name server. This is useful when running multiple services (like an FTP and a webserver) from a single IP address. Each service can then have its own entry in DNS (like ftp.example.com and www.example.com.)
- An MX record or mail exchange record maps a domain name to a list of mail exchange servers for that domain.
- A PTR record or pointer record maps an IPv4 address to the canonical name for that host. Setting up a PTR record for a hostname in the in-addr.arpa domain that corresponds to an IP address implements reverse DNS lookup for that address. For example (at the time of writing), www.icann.net has the IP address 192.0.34.164, but a PTR record maps 164.34.0.192.in-addr.arpa to its canonical name, referrals.icann.org.
- An NS record or name server record maps a domain name to a list of DNS servers authoritative for that domain. Delegations depend on NS records.
- An SOA record or start of authority record specifies the DNS server providing authoritative information about an Internet domain, the email of the domain administrator, the domain serial number, and several timers relating to refreshing the zone.
- An SRV record is a generalized service location record.
- A TXT record allows an administrator to insert arbitrary text into a DNS record. For example, this record is used to implement the Sender Policy Framework and DomainKeys specifications.
- NAPTR records ("Naming Authority Pointer") are a newer type of DNS record that support regular expression based rewriting.

Other types of records simply provide information (for example, a LOC record gives the physical location of a host), or experimental data (for example, a WKS record gives a list of servers offering some well known service such as HTTP or POP3 for a domain).

Internationalised Domain Names

While domain names in the DNS have no restrictions on the characters they use and can include non-ASCII characters, the same is not true for host names. Host names are the names most people see and use for things like e-mail and web browsing. Host names are restricted to a small subset of the ASCII character set that includes the Roman alphabet in upper and lower case, the digits 0 through 9, the dot, and the hyphen. (See RFC 3696 section 2 for details.) This prevented the representation of names and words of many languages natively. ICANN has approved the Punycode-based IDNA system, which maps Unicode strings into the valid DNS character set, as a workaround to this issue. Some registries have adopted IDNA.

Security Issues In DNS

DNS was not originally designed with security in mind, and thus has a number of security issues. DNS responses are traditionally not cryptographically signed, leading to many attack possibilities; DNSSEC modifies DNS to add support for cryptographically signed responses. There are various extensions to support securing zone transfer information as well.

Some domain names can spoof other, similar-looking domain names. For example, "paypal.com" and "paypa1.com" are different names, yet users may be unable to tell the difference. This problem is much more serious in systems that support internationalized domain names, since many characters that are different (from the point of view of ISO 10646) appear identical on typical computer screens.

Legal Users Of Domains

Registrant

No one in the world really "owns" a domain name except the Network Information Centre (NIC), or domain name registry. Most of the NICs in the world receive an annual fee from a legal user in order for the legal user to utilize the domain name (i.e. a sort of a leasing agreement exists, subject to the registry's terms and conditions). Depending on the various naming convention of the registries, legal users become commonly known as "registrants" or as "domain holders".

ICANN holds a complete list of domain registries in the world. One can find the legal user of a domain name by looking in the WHOIS database held by most domain registries.

For most of the more than 240 country code top-level domains (ccTLDs), the domain registries hold the authoritative WHOIS (Registrant, name servers, expiry dates etc.) For instance, DENIC, Germany NIC holds the authoritative WHOIS to a .DE domain name.

However, some domain registries, such as for .COM, .ORG, .INFO, etc., use a registry-registrar model. There are hundreds of Domain Name Registrars that actually perform the domain name registration with the end-user (see lists at ICANN or VeriSign). By using this method of distribution, the registry only has to manage the relationship with the registrar, and the registrar maintains the relationship with the end-users, or 'registrants'. For .COM, .NET domain names, the domain registries, VeriSign holds a basic WHOIS (registrar and name servers etc.) One can find the detailed WHOIS (Registrant, name servers, expiry dates etc.) at the registrars.

Since about 2001, most gTLD registries (.ORG, .BIZ, .INFO) have adopted a so-called "thick" registry approach, i.e. keeping the authoritative WHOIS with the various registries instead of the registrars.

Administrative Contact

A registrant usually designates an administrative contact to manage the domain name. In practice, the administrative contact usually has the most immediate power over a domain. Management functions delegated to the administrative contacts may include (for example):

- the obligation to conform to the requirements of the domain registry in order to retain the right to use a domain name
- authorization to update the physical address, e-mail address and telephone number etc. in WHOIS

Technical Contact

A technical contact manages the name servers of a domain name. The many functions of a technical contact include:

- making sure the configurations of the domain name conforms to the requirements of the domain registry
- updating the domain zone
- providing the 24x7 functionality of the name servers (that leads to the accessibility of the domain name)

Billing Contact

The party whom a NIC invoices.

Name Servers

Namely the authoritative name servers that host the domain name zone of a domain name.

Politics

Many investigators have voiced criticism of the methods currently used to control ownership of domains. Critics commonly claim abuse by monopolies or near-monopolies, such as VeriSign, Inc. Particularly noteworthy was the VeriSign Site Finder system which redirected all unregistered .com and .net domains to a VeriSign webpage. Despite widespread criticism, VeriSign only reluctantly removed it after ICANN threatened to revoke its contract to administer the root name servers.

There is also significant disquiet regarding United States political influence over the Internet Corporation for Assigned Names and Numbers (ICANN). This was a significant issue in the attempt to create a .xxx Top-level domain and sparked greater interest in Alternative DNS roots that would be beyond the control of any single country.

Truth in Domain Names Act

In the United States, the "Truth in Domain Names Act", in combination with the PROTECT Act, forbids the use of a misleading domain name with the intention of attracting people into viewing a visual depiction of sexually explicit conduct on the Internet.

The Internet – Websites

What Is It?

A website (or Web site) is a collection of web pages, images, videos and other digital assets and hosted on a particular domain or subdomain on the World Wide Web.

A web page is a document, typically written in HTML, that is almost always accessible via HTTP, a protocol that transfers information from the website's server to display in the user's web browser.

All publicly accessible websites are seen collectively as constituting the "World Wide Web".

The pages of websites can usually be accessed from a common root URL called the homepage, and usually reside on the same physical server. The URLs of the pages organize them into a hierarchy, although the hyperlinks between them control how the reader perceives the overall structure and how the traffic flows between the different parts of the sites.

Some websites require a subscription to access some or all of their content. Examples of subscription sites include many business sites, parts of many news sites, gaming sites, message boards, Web-based e-mail services, and sites providing real-time stock market data.

History

The first on-line website appeared in 1991. On 30 April 1993, CERN announced that the World Wide Web would be free to anyone. A copy of the original first Web page, created by Tim Berners-Lee, is kept @
http://www.w3.org/History/19921103-hypertext/hypertext/WWW/TheProject.html

Overview

Organized by function a website may be

- a personal website
- a business website
- a government website or
- a non-profit organization website.

It could be the work of an individual, a business or other organization and is typically dedicated to some particular topic or purpose. Any website can contain a hyperlink to any other website, so the distinction between individual sites, as perceived by the user, may sometimes be blurred.

Websites are written in, or dynamically converted to, HTML (Hyper Text Markup Language) and are accessed using a software program called a Web browser, also known as an HTTP client. Web pages can be viewed or otherwise accessed from a range of computer based and Internet enabled devices of various sizes, including desktop computers, laptop computers, PDAs and cell phones.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

A website is hosted on a computer system known as a web server, also called an HTTP server, and these terms can also refer to the software that runs on these system and that retrieves and delivers the Web pages in response to requests from the website users. Apache is the most commonly used Web server software (according to Netcraft statistics) and Microsoft's Internet Information Server (IIS) is also commonly used.

A static website, is one that has web pages stored on the server in the same form as the user will view them. They are edited using three broad categories of software:

- Text editors. such as Notepad or TextEdit, where the HTML is manipulated directly within the editor program
- WYSIWYG editors. such as Microsoft FrontPage and Macromedia Dreamweaver, where the site is edited using a GUI interface and the underlying HTML is generated automatically by the editor software
- Template-based editors, such as Rapidweaver and iWeb, which allow users to quickly create and upload websites to a web server without having to know anything about HTML, as they just pick a suitable template from a palette and add pictures and text to it in a DTP-like fashion without ever having to see any HTML code.

A dynamic website is one that has frequently changing information or collates information on the hop each time a page is requested. For example, it would call various bits of information from a database and put them together in a pre-defined format to present the reader with a coherent page. It interacts with users in a variety of ways including by reading cookies recognizing users' previous history, session variables, server side variables etc., or by using direct interaction (form elements, mouseovers, etc.). A A site can display the current state of a dialogue between users, monitor a changing situation, or provide information in some way personalized to the requirements of the individual user.

There is a wide range of software systems, such as ColdFusion (CFM), Active Server Pages (ASP), Java Server Pages (JSP) and the PHP programming language that are available to generate dynamic Web systems and dynamic sites. Sites may also include content that is retrieved from one or more databases or by using XML-based technologies such as RSS.

Static content may also be dynamically generated either periodically, or if certain conditions for regeneration occur (cached) in order to avoid the performance loss of initiating the dynamic engine on a per-user or per-connection basis.

Plugins are available to expand the features and abilities of Web browsers, which use them to show active content, such as Flash, Shockwave or applets written in Java. Dynamic HTML also provides for user interactivity and realtime element updating within Web pages (i.e., pages don't have to be loaded or reloaded to effect any changes), mainly using the DOM and JavaScript, support which is built-in to most modern Web browsers.

Websites As Businesses

While most business websites serve as a shop window for brick and mortar businesses it is increasingly the case that some websites are businesses in their own right. These websites are fully self-contained businesses entities offering, for example, immediate downloads of commercial software on payment of the product's price via their shopping cart. Others may offer no products at all but provide free information with income coming from clicks the visitors make on advertisements (see contextual ads). There is a wide range of monetizing used on such sites and the sites themselves are actively traded and bought and sold as going concerns.

Spelling

As noted above, there are several different spellings for this term. Although "website" is used commonly, especially in British English, the Associated Press Stylebook, Reuters, Microsoft, academia, book publishing, The Chicago Manual of Style, and dictionaries such as Merriam-Webster use the two-word, initially capitalized spelling Web site. This is because "Web" is not a general term but a shortened form of World Wide Web. As with many newly created terms, it may take some time before a common spelling is finalized. (This controversy also applies to derivative terms such as "Web master"/"webmaster" and "Web cam"/"webcam").

The Canadian Oxford Dictionary and the Canadian Press Stylebook list "website" and "web page" as the preferred spellings. The Oxford Dictionary began using "website" as its standardized form in 2004.

Bill Walsh, the copy chief of The Washington Post's national desk, and one of American English's foremost grammarians, argues for the two-word spelling in his books Lapsing into a Comma and The Elephants of Style, and on his site.

Types Of Websites

There are many varieties of Web sites, each specialising in a particular type of content or use, and they may be arbitrarily classified in any number of ways. A few such classifications might include:

- Affiliate: enabled portal that renders not only its custom CMS but also syndicated content from other content providers for an agreed fee. There are usually three relationship tiers. Affiliate Agencies (e.g., Commission Junction), Advertisers (e.g., Ebay) and consumer (e.g., Yahoo).
- Archive site: used to preserve valuable electronic content threatened with extinction. Two examples are: Internet Archive, which since 1996 has preserved billions of old (and new) Web pages; and Google Groups, which in early 2005 was archiving over 845,000,000 messages posted to Usenet news/discussion groups.
- Blog (or web log) site: sites generally used to post online diaries which may include discussion forums (e.g., blogger, Xanga).
- Corporate website: used to provide background information about a business, organization, or service.
- Commerce site or eCommerce site: for purchasing goods, such as Amazon.com.
- Community site: a site where persons with similar interests communicate with each other, usually by chat or message boards, such as MySpace.
- Database site: a site whose main use is the search and display of a specific database's content such as the Internet Movie Database or the Political Graveyard.
- Development site: a site whose purpose is to provide information and resources related to software development, Web design and the like.
- Directory site: a site that contains varied contents which are divided into categories and subcategories, such as Yahoo! directory, Google directory and Open Directory Project.
- Download site: strictly used for downloading electronic content, such as software, game demos or computer wallpaper.
- Employment site: allows employers to post job requirements for a position or positions and prospective employees to fill an application.
- Game site: a site that is itself a game or "playground" where many people come to play, such as MSN Games and Pogo.com.

- Geodomain refers to domain names that are the same as those of geographic entities, such as cities and countries. For example, Richmond.com is the geodomain for Richmond, Virginia.
- Gripe site: a site devoted to the critique of a person, place, corporation, government, or institution.
- Humor site: satirizes, parodies or otherwise exists solely to amuse.
- Information site: contains content that is intended to inform visitors, but not necessarily for commercial purposes, such as: RateMyProfessors.com, Free Internet Lexicon and Encyclopedia. Most government, educational and non-profit institutions have an informational site.
- Java applet site: contains software to run over the Web as a Web application.
- Mirror (computing) site: A complete reproduction of a website.
- News site: similar to an information site, but dedicated to dispensing news and commentary.
- Personal homepage: run by an individual or a small group (such as a family) that contains information or any content that the individual wishes to include.
- Phish site: a website created to fraudulently acquire sensitive information, such as passwords and credit card details, by masquerading as a trustworthy person or business (such as Social Security Administration, PayPal) in an electronic communication. (see Phishing).
- Political site: A site on which people may voice political views.
- Pornography (porn) site: a site that shows pornographic images and videos.
- Rating site: A site on which people can praise or disparage what is featured (e.g. ratemycar.com).
- Review site: A site on which people can post reviews for products or services.
- Search engine site: a site that provides general information and is intended as a gateway or lookup for other sites. A pure example is Google, and the most widely known extended type is Yahoo!.
- Shock site: includes images or other material that is intended to be offensive to most viewers (e.g. rotten.com).
- Web portal: a site that provides a starting point or a gateway to other resources on the Internet or an intranet.
- Wedsite: a website that details a couple's wedding event, often sharing stories, photos, and event information.
- Wiki site: a site which users collaboratively edit (such as Wikipedia).

Some websites may be included in one or more of these categories. For example, a business website may promote the business's products, but may also host informative documents, such as white papers. There are also numerous sub-categories to the ones listed above. For example, a porn site is a specific type of eCommerce site or business site (that is, it is trying to sell memberships for access to its site). A fan site may be a dedication from the owner to a particular celebrity.

Websites are constrained by architectural limits (e.g., the computing power dedicated to the website). Very large websites, such as Yahoo!, Microsoft, and Google employ many servers and load balancing equipment such as Cisco Content Services Switches to distribute visitor loads over multiple computers at multiple locations.

In January of 2007, Netcraft, an Internet monitoring company that has tracked Web growth since 1995, reported that there were 106,875,138 Web sites with domain names and content on them in 2007, compared to just 18,000 Web sites in August 1995.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

The Internet – Meta Elements

What Is It?

Meta elements are HTML elements used to provide structured metadata about a web page. Such elements must be placed as tags in the head section of an HTML document.

An HTML element indicates structure in an HTML document and a way of hierarchically arranging content. More specifically, an HTML element is an SGML element that meets the requirements of one or more of the HTML Document Type Definitions (DTDs). These elements have properties: both attributes and content, as specified (both allowable and required) according to the appropriate HTML DTD (for example, the HTML 4.01 strict DTD). Elements may represent headings, paragraphs, hypertext links, lists, embedded media, and a variety of other structures.

Metadata has multiple definitions, the briefest of which is "data about data." It can generally be thought of as information that describes, or supplements, the central data. For example, metadata produced by digital still cameras describe the settings used for the picture, such as exposure value or flash intensity. In such cases, the metadata can be considered as extra data, which merely add information, and is not critical to the functions of the main data. In other cases, such as on a Zip disk, metadata provides information regarding the write-protected status of the disk. In such a case, then, metadata can be considered essential to the proper functioning of the main product. Since the difference between data and metadata is often subtle and context-specific, rigorous definitions are elusive.

Even when it is not essential to the proper functioning of a product, metadata is valuable because of the context that it provides, and the ways that contextual information can be used. When data is made available to a potential user, the user (human or computer) must put the data into an existing model of knowledge, and may ask questions to do so. For example, in the case of an image, typical questions include "When was this taken?" and "Who and what are in this image?" Metadata provides context to answer many of these questions. In sophisticated data systems, the metadata -- the contextual information surrounding the data -- will also be very sophisticated, capable of answering many questions that help understand the data.

Meta Tag Use In Search Engine Optimization

Meta elements provide information about a given webpage, most often to help search engines categorize them correctly, and are inserted into the HTML code in the format illustrated above, but are not visible to a user looking at the site.

They have been the focus of a field of marketing research known as search engine optimization (SEO), where different methods are explored to provide a user's site with a higher ranking on search engines. In the mid to late 1990s, search engines were reliant on meta data to correctly classify a web page and webmasters quickly learned the commercial significance of having the right meta element, as it frequently led to a high ranking in the search engines - and thus, high traffic to the web site.

As search engine traffic achieved greater significance in online marketing plans, consultants were brought in who were well versed in how search engines perceive a web site. These consultants used a variety of techniques (legitimate and otherwise) to improve ranking for their clients.

The keyword tag was popularized by search engines such as Infoseek and AltaVista in 1996 and its popularity quickly grew until it became one of the most commonly used META tags. By late 1997, however, search engine providers realised that information stored in META tags, especially the keyword tag, could be unreliable and misleading, and at worst, could be used to draw users into spam sites. (Unscrupulous webmasters could easily place false keywords into a META tag in order to draw people to their site, whether the content matched these keywords or not.)

Search engines began dropping support for META keywords in 1998, and by the early 2000s, most search engines had veered away from reliance on meta elements, and in July 2002 AltaVista, one of the last major search engines to still offer support, finally stopped including them.

Newer search engines like Google and FAST have never had any support for the META keywords tag.

Some academics, e.g. Zhang & Dimitroff, 2004, have concluded that "webpages with metadata elements achieved better visibility performance than those without metadata elements". However, this may be due to confusion between the HTML <title> tag, which Google does use (and is very important), and HTML <meta...> tags, which Google does not use for indexing. The Director of Research at Google, Monika Henziger, was quoted (in 2002) as saying, "Currently we don't trust metadata". Techniques have also been developed in order to penalize web sites considered to be "cheating the system". For example, a web site repeating the same meta keyword several times may have its ranking decreased by a search engine trying to eliminate this practice, though that is unlikely. It's more likely that a search engine will ignore the meta keyword element completely, and most do regardless of how many words used in the element.

Some search engines such as Google will display the text specified in the content of the META description tag for a page in their result listings. This allows the webpage author to give a more meaningful description for listings than might be displayed if the search engine was to automatically create its own description based on the page content.

Meta Tag Use In Social Bookmarking

In contrast to completely automated systems like search engines, author-supplied metadata can be useful in situations where the page content has been vetted as trustworthy by a reader. An example usage would be a social bookmarking site that supplies the meta description for the notes and the meta author as a tag. Since the reader is bookmarking the page, it can be assumed he trusts it to some degree. The meta tags act as a convenience for the user.

Redirects

Meta refresh elements can be used to instruct a web browser to automatically refresh a web page after a given time interval. It is also possible to specify an alternative URL and use this technique in order to redirect the user to a different location. Using a meta refresh in this way and solely by itself will not always work. For Internet Explorer's security settings, under the miscellaneous category, meta refresh can be turned off by the user, therby disabling its redirect ability entirely.

Many web design tutorials also point out that client side redirecting tends to interfere with the normal functioning of a web browser's "back" button. After being redirected, clicking the back button will cause the user to go back to the redirect page, which redirects them again. Some modern browsers seem to overcome this problem, however, including Safari, Mozilla Firefox and Opera.

HTTP Message Headers

Meta tags of the form <meta http-equiv="foo" content="bar"> can be used to send http headers. For example,

<meta http-equiv="Expires" content="Wed, 21 Jun 2006 14:25:27 GMT">

would tell a browser (or other http client) that the page "expires" on June 21, 2006 and that it may safely cache the page until then.

Alternative To META Elements

An alternative to META elements for enhanced subject access within a web site is the use of a back-of-book-style index for the web site. See examples at the web sites of the Australian Society of Indexers www.aussi.org and the American Society of Indexers (Journal of Internet Cataloging, Volume 5(1), 2002).

In 1994, ALIWEB, which was likely the first web search engine, also used an index file to provide the type of information commonly found in meta keywords tags.

The Internet – Search Engines

What Is It?

A search engine is an information retrieval system designed to help find information stored on a computer system, such as on the World Wide Web, inside a corporate or proprietary network, or in a personal computer. The search engine allows one to ask for content meeting specific criteria (typically those containing a given word or phrase) and retrieves a list of items that match those criteria. This list is often sorted with respect to some measure of relevance of the results. Search engines use regularly updated indexes to operate quickly and efficiently.

Without further qualification, search engine usually refers to a Web search engine, which searches for information on the public Web. Other kinds of search engine are enterprise search engines, which search on intranets, personal search engines, and mobile search engines. Different selection and relevance criteria may apply in different environments, or for different uses.

Some search engines also mine data available in newsgroups, databases, or open directories. Unlike Web directories, which are maintained by human editors, search engines operate algorithmically or are a mixture of algorthmic and human input.

History

The very first tool used for searching on the Internet was Archie. The name stands for "archive" without the "v". It was created in 1990 by Alan Emtage, a student at McGill University in Montreal. The program downloaded the directory listings of all the files located on public anonymous FTP (File Transfer Protocol) sites, creating a searchable database of filenames; however, Archie could not search by file contents.

While Archie indexed computer files, Gopher indexed plain text documents. Gopher was created in 1991 by Mark McCahill at the University of Minnesota: Gopher was named after the school's mascot.[1] Because these were text files, most of the Gopher sites became websites after the creation of the World Wide Web.

Two other programs, Veronica and Jughead, searched the files stored in Gopher index systems. Veronica (Very Easy Rodent-Oriented Net-wide Index to Computerized Archives) provided a keyword search of most Gopher menu titles in the entire Gopher listings. Jughead (Jonzy's Universal Gopher Hierarchy Excavation And Display) was a tool for obtaining menu information from various Gopher servers. While the name of the search engine "Archie" was not a reference to the Archie comic book series, "Veronica" and "Jughead" are characters in the series, thus referencing their predecessor.

The first Web search engine was Wandex, a now-defunct index collected by the World Wide Web Wanderer, a web crawler developed by Matthew Gray at MIT in 1993. Another very early search engine, Aliweb, also appeared in 1993, and still runs today. The first "full text" crawler-based search engine was WebCrawler, which came out in 1994. Unlike its predecessors, it let users search for any word in any webpage, which became the standard for all major search engines since. It was also the first one to be widely known by the public. Also in 1994 Lycos (which started at Carnegie Mellon University) came out, and became a major commercial endeavor.

Soon after, many search engines appeared and vied for popularity. These included Excite, Infoseek, Inktomi, Northern Light, and AltaVista. In some ways, they competed with popular directories such as Yahoo!. Later, the directories integrated or added on search engine technology for greater functionality.

Search engines were also known as some of the brightest stars in the Internet investing frenzy that occurred in the late 1990s. Several companies entered the market spectacularly, receiving record gains during their initial public offerings. Some have taken down their public search engine, and are marketing enterprise-only editions, such as Northern Light.

Google

Around 2001, the Google search engine rose to prominence. Its success was based in part on the concept of link popularity and PageRank. The number of other websites and webpages that link to a given page is taken into consideration with PageRank, on the premise that good or desirable pages are linked to more than others. The PageRank of linking pages and the number of links on these pages contribute to the PageRank of the linked page. This makes it possible for Google to order its results by how many websites link to each found page. Google's minimalist user interface is very popular with users, and has since spawned a number of imitators.

Google and most other web engines utilize not only PageRank but more than 150 criteria to determine relevancy. The algorithm "remembers" where it has been and indexes the number of cross-links and relates these into groupings. PageRank is based on citation analysis that was developed in the 1950s by Eugene Garfield at the University of Pennsylvania. Google's founders cite Garfield's work in their original paper. In this way virtual communities of webpages are found. Teoma's search technology uses a communities approach in its ranking algorithm. NEC Research Institute has worked on similar technology. Web link analysis was first developed by Jon Kleinberg and his team while working on the CLEVER project at IBM's Almaden Research Center. Google is currently the most popular search engine.

Baidu

Due to the difference between hanzi and the Roman alphabet, the Chinese search market did not boom until the introduction of Baidu in 2000.

Yahoo! Search

The two founders of Yahoo!, David Filo and Jerry Yang, Ph.D. candidates in Electrical Engineering at Stanford University, started their guide in a campus trailer in February 1994 as a way to keep track of their personal interests on the Internet. Before long they were spending more time on their home-brewed lists of favourite links than on their doctoral dissertations. Eventually, Jerry and David's lists became too long and unwieldy, and they broke them out into categories. When the categories became too full, they developed subcategories ... and the core concept behind Yahoo! was born. In 2002, Yahoo! acquired Inktomi and in 2003, Yahoo! acquired Overture, which owned AlltheWeb and AltaVista. Despite owning its own search engine, Yahoo! initially kept using Google to provide its users with search results on its main website Yahoo.com. However, in 2004, Yahoo! launched its own search engine based on the combined technologies of its acquisitions and providing a service that gave pre-eminence to the Web search engine over the directory.

Ask.com

Ask.com, or "Ask Jeeves" claims to be the first commercial question-answering search engine for the World Wide Web. It is owned by IAC. In 2006, it became the 4th largest search engine in world.

Microsoft

The most recent major search engine is MSN Search (evolved into Windows Live Search), owned by Microsoft, which previously relied on others for its search engine listings. In 2004 it debuted a beta version of its own results, powered by its own web crawler (called msnbot). In early 2005 it started showing its own results live. This was barely noticed by average users unaware of where results come from, but was a huge development for many webmasters, who seek inclusion in the major search engines. At the same time, Microsoft ceased using results from Inktomi, now owned by Yahoo!. In 2006, Microsoft migrated to a new search platform - Windows Live Search, retiring the "MSN Search" name in the process.

Challenges Faced By Search Engines

- The Web is growing much faster than any present-technology search engine can possibly index (see distributed web crawling). In 2006, some users found major search-engines became slower to index new webpages.
- Many webpages are updated frequently, which forces the search engine to revisit them periodically.
- The queries one can make are currently limited to searching for key words, which may result in many false positives, especially using the default whole-page search. Better results might be achieved by using a proximity-search option with a search-bracket to limit matches within a paragraph or phrase, rather than matching random words scattered across large pages. Another alternative is using human operators to do the researching for the user with organic search engines.
- Dynamically generated sites may be slow or difficult to index, or may result in excessive results, perhaps generating 500 times more webpages than average. Example: for a dynamic webpage which changes content based on entries inserted from a database, a search-engine might be requested to index 50,000 static webpages for 50,000 different parameter values passed to that dynamic webpage.
- Many dynamically generated websites are not indexable by search engines; this phenomenon is known as the invisible web. There are search engines that specialize in crawling the invisible web by crawling sites that have dynamic content, require forms to be filled out, or are password protected.
- Relevancy: sometimes the engine can't get what the person is looking for.
- Some search-engines do not rank results by relevance, but by the amount of money the matching websites pay.
- In 2006, hundreds of generated websites used tricks to manipulate a search-engine to display them in the higher results for numerous keywords. This can lead to some search results being polluted with linkspam or bait-and-switch pages which contain little or no information about the matching phrases. The more relevant webpages are pushed further down in the results list, perhaps by 500 entries or more.
- Secure pages (content hosted on HTTPS URLs) pose a challenger for crawlers which either can't browse the content for technical reasons or won't index it for privacy reasons.

How Search Engines Work

A search engine operates, in the following order

1. Web crawling
2. Indexing
3. Searching

Web search engines work by storing information about a large number of web pages, which they retrieve from the WWW itself. These pages are retrieved by a Web crawler (sometimes also known as a spider) — an automated Web browser which follows every link it sees. Exclusions can be made by the use of robots.txt. The contents of each page are then analyzed to determine how it should be indexed (for example, words are extracted from the titles, headings, or special fields called meta tags). Data about web pages are stored in an index database for use in later queries. Some search engines, such as Google, store all or part of the source page (referred to as a cache) as well as information about the web pages, whereas others, such as AltaVista, store every word of every page they find. This cached page always holds the actual search text since it is the one that was actually indexed, so it can be very useful when the content of the current page has been updated and the search terms are no longer in it. This problem might be considered to be a mild form of linkrot, and Google's handling of it increases usability by satisfying user expectations that the search terms will be on the returned webpage. This satisfies the principle of least astonishment since the user normally expects the search terms to be on the returned pages. Increased search relevance makes these cached pages very useful, even beyond the fact that they may contain data that may no longer be available elsewhere.

When a user comes to the search engine and makes a query, typically by giving key words, the engine looks up the index and provides a listing of best-matching web pages according to its criteria, usually with a short summary containing the document's title and sometimes parts of the text. Most search engines support the use of the boolean terms AND, OR and NOT to further specify the search query. An advanced feature is proximity search, which allows users to define the distance between keywords.

The usefulness of a search engine depends on the relevance of the result set it gives back. While there may be millions of webpages that include a particular word or phrase, some pages may be more relevant, popular, or authoritative than others. Most search engines employ methods to rank the results to provide the "best" results first. How a search engine decides which pages are the best matches, and what order the results should be shown in, varies widely from one engine to another. The methods also change over time as Internet usage changes and new techniques evolve.

Most Web search engines are commercial ventures supported by advertising revenue and, as a result, some employ the controversial practice of allowing advertisers to pay money to have their listings ranked higher in search results. Those search engines which do not accept money for their search engine results make money by running search related ads alongside the regular search engine results. The search engines make money everytime someone clicks on one of these ads.

The vast majority of search engines are run by private companies using proprietary algorithms and closed databases, though some are open source.

Web Crawling

A web crawler (also known as a Web spider or Web robot) is a program or automated script which browses the World Wide Web in a methodical, automated manner. Other less frequently used names for Web crawlers are ants, automatic indexers, bots, and worms (Kobayashi and Takeda, 2000).

This process is called Web crawling or spidering. Many sites, in particular search engines, use spidering as a means of providing up-to-date data. Web crawlers are mainly used to create a copy of all the visited pages for later processing by a search engine, that will index the downloaded pages to provide fast searches. Crawlers can also be used for automating maintenance tasks on a Web site, such as checking links or validating HTML code. Also, crawlers can be used to gather specific types of information from Web pages, such as harvesting e-mail addresses (usually for spam).

A Web crawler is one type of bot, or software agent. In general, it starts with a list of URLs to visit, called the seeds. As the crawler visits these URLs, it identifies all the hyperlinks in the page and adds them to the list of URLs to visit, called the crawl frontier. URLs from the frontier are recursively visited according to a set of policies.

Indexing

Search engine indexing entails how data is collected, parsed, and stored to facilitate fast and accurate retrieval. Index design incorporates interdisciplinary concepts from Linguistics, Cognitive psychology, Mathematics, Informatics, Physics, and Computer science. An alternate name for the process is Web indexing, within the context of search engines designed to find web pages on the Internet.

Popular engines focus on the full-text indexing of online, natural language documents, yet there are other searchable media types such as video, audio, and graphics. Meta search engines reuse the indices of other services and do not store a local index, whereas cache-based search engines permanently store the index along with the corpus. Unlike full text indices, partial text services restrict the depth indexed to reduce index size. Larger services typically perform indexing at a predetermined interval due to the required time and processing costs, whereas agent-based search engines index in real time.

Storage Costs & Crawling Time

Storage costs are not the limiting resource in search engine implementation. Simply storing 10 billion pages of 10 kbytes each (compressed) requires 100TB and another 100TB or so for indexes, giving a total hardware cost of under $200k: 100 cheap PCs each with four 500GB disk drives.

However, a public search engine requires considerably more resources than this to calculate query results and to provide high availability. Also, the costs of operating a large server farm are not trivial.

Crawling 10B pages with 100 machines crawling at 100 pages/second would take 1M seconds, or 11.6 days on a very high capacity Internet connection. Most search engines crawl a small fraction of the Web (10-20% pages) at around this frequency or better, but also crawl dynamic websites (e.g. news sites and blogs) at a much higher frequency.

Local Search

On the World Wide Web, local search is a query that uses a special type of search engine. The query includes not only information about "what" the site visitor is searching for (such as keywords, a business category, or the name of a consumer product) but also "where" information, such as a street address, city name, postal code, or geographic coordinates like latitude and longitude. These sites are primarily supported by advertising, but some accept payments from retail merchants and others wishing to be listed.

Various geolocation techniques may be used to match visitors' queries with information of interest. The type of information and points of interest returned varies with the type of local search engine. Some sites, such as superpages.com (owned by Idearc Media) and TrueLocal, list local businesses, usually in a categorization scheme similar to (or actually derived from) a traditional Yellow Pages phone book.

Google Maps (formerly Google Local) looks for physical addresses mentioned in regular web pages. It provides these results to visitors, along with business listings and maps. Product-specific search engines, such as Froogle and Yokel.com use techniques such as targeted web crawling and direct feeds to collect information about products for sale in a specific geographic area. Other local search engines adjunct to major web search portals include general Windows Live Local, Yahoo! Local, and ask.com's AskCity. Other popular local-only search engines include Local.com.

Traditional media companies, including newspaper publishers and television and radio broadcasters, are starting to add local search to their local websites in an effort to attract their share of local search traffic and advertising revenues in the markets they serve. These local media companies either develop their own technology and data, or purchase "private label" services from a local search company.

Electronic publishers (such as businesses or individuals) with information they would like to appear on local search engines have several options. Business listing information can be distributed via the traditional Yellow Pages, electronic Yellow Pages aggregators, direct contact with the local search engines (such as through Google Base), or search engine optimization services. Some search engines will pick up on web pages that contain regular street addresses displayed in machine-readable text (rather than a picture of text, which is more difficult to interpret). Web pages can also use GeoTagging techniques, such as microformats.

Types Of Search Engine

- General Search Engines
- Open Source Search Engines
- Metasearch Engines
- Regional Search Engines
- People Search Engines
- Email-Based Search Engines
- Visual Search Engines
- Answer-Based Search Engines
- Google-Based Search Engines
- Yahoo!-Based Search Engines
- ODP Based Search Engines
- Windows-Live-Based Search Engines
- Job Search Engines
- Blog Search Engines
- News Search Engines
- Multimedia Search Engines
- Code Search Engines
- Bittorrent Search Engines

- Accountancy Search Engines
- Medical Search Engines
- Property Search Engines
- Business Search Engines
- Comparison Shopping Search Engines
- Charity Search Engines
- Geographic Search Engines
- Social Search Engines
- Search Engines For Kids
- Desktop Search Engines
- Defunct Search Engines

A List Of Some Search Engines By Type

General search engines

- Alexa Internet
- Alta Vista
- Ask.com (formerly Ask Jeeves)
- AskMeNow
- Baidu
- Dogpile
- Exalead
- Gigablast
- Google
- MozDex
- Windows Live Search (formerly MSN Search)
- WiseNut
- Yahoo! Search

Open source search engines

- DataparkSearch
- Egothor
- Gonzui
- Ht://dig
- Lucene
- mnoGoSearch
- Namazu
- Nutch
- OpenFTS
- SWISH-E
- Wikiasari
- Xapian
- YaCy
- Zettair

Metasearch engines

- Bioinformatic_Harvester
- Brainboost
- Clusty
- Dogpile

- Excite
- HotBot
- Info.com
- Ixquick
- Kartoo
- Mamma
- Metacrawler
- MetaLib
- Myriad Search
- SideStep
- WebCrawler
- Fazzle.com

Regional search engines

- Accoona, China/US
- Alleba, Philippines
- Ansearch, Australia/US/UK/NZ
- Araby, Middle East
- Baidu, China
- Daum, Korea
- In.gr, Greece
- Jabse, Bulgaria
- Naver, Korea
- Rambler, Russia
- Rediff, India
- SAPO, Portugal
- Yandex, Russia

People search engines

- Zoominfo

Email-based search engines

- TEK

Visual search engines

- Grokker
- Kartoo
- Quintura

Answer-based search engines

- Answers
- AskMeNow
- BrainBoost
- Lexxe
- Lycos iQ
- Windows Live QnA
- Yahoo! Answers

Google-based search engines

- AOL Search
- Netscape
- Google

Yahoo!-based search engines

- AltaVista
- AlltheWeb
- GoodSearch
- Yahoo

ODP based search engines

- Zitku

Windows-Live-based search engines

- A9.com
- Lycos

Job search engines

- Naukri.com (India)
- Bixee.com (India)
- Craigslist (by city)
- Eluta.ca (Canada)
- CareerBuilder.com (USA)
- Hotjobs.com (USA)
- Indeed.com (USA)
- Monster.com (USA)
- Recruit.net (International)
- SimplyHired.com (USA)

Blog search engines

- Bloglines
- IceRocket
- PubSub
- Sphere
- Technorati

News search engines

- Google News
- MagPortal
- Newslookup
- Topix.net
- Yahoo! News

Multimedia search engines

- Picsearch
- Podscope
- Singingfish
- blinkx

Code search engines

- Krugle
- Koders
- Google Code Search

BitTorrent search engines

- BitTorrent
- Isohunt
- Mininova
- The Pirate Bay
- TorrentSpy
- FlixFlux

Accountancy search engines

- IFACnet

Medical search engines

- Entrez (includes Pubmed)
- GoPubMed (knowledge-based: GO - GeneOntology)
- KMLE Medical Dictionary
- MeshPubMed (knowledge-based: MeSH - Medical Subject Headings)
- WebMD

Property search engines

- Home.co.uk
- Rightmove
- Zillow.com

Business search engines

- Thomasnet (United States)
- Business.com
- American Indian Search

Comparison shopping search engines

- Froogle
- Kelkoo
- MSN Shopping
- MySimon
- NexTag
- PriceRunner
- Shopzilla

- Retrevo

Charity search engines

- GoodSearch

Geographic search engines

- Géoportail
- Google Maps
- MapQuest
- Virtual Earth
- Yahoo! Maps

Social search engines

- Google Coop is a platform provided by Google that allows web developers to feature specialized information in web searches, refine and categorise queries, and create customized search engines
- Rollyo
- Wink provides web search by analyzing user contributions such as bookmarks and feedback

Search Engines for Kids

- Yahoo! Kids
- Ask for Kids
- Quintura for Kids

Desktop search engines

- Ask.com
- Autonomy — IDOL Enterprise Desktop Search.
- Beagle — Open source desktop search tool for Linux based on Lucene
- Blinkx
- Copernic Desktop Search
- Docco — based on Apache's indexing and search engine Lucene, and it requires a Java Runtime Environment.
- exalead one:desktop
- Fast
- Filehawk — Indexes and searches files' content in computers, networks and removable data storage devices
- Gaviri PocketSearch — Indexes desktop and portable devices.
- Google Desktop — Integrates with the main Google search engine page.
- HotBot — Lycos HotBot has an adware desktop search toolbar for IE.
- imgSeek — Desktop content-based image search
- ISYS Search Software — ISYS:desktop search software.
- Likasoft Archivarius 3000
- Windows Desktop Search — MSN Search Toolbar includes Windows Desktop Search which incorporates much of the desktop search technology initially promised for Windows Vista, the next version of Microsoft Windows. The search integrates into the task bar and Internet Explorer windows.
- Spotlight — Found in Apple Mac OS X "Tiger".

- Tropes Zoom — Semantic Search Engine.
- X1 Desktop Search, renamed X1 Enterprise Client

Defunct search engines

- Grub
- Infoseek
- Inktomi
- Northern Light
- Overture.com (formerly GoTo.com)
- Teoma
- World Wide Web Worm

The Internet – E-Mail

What Is It?

Electronic mail (abbreviated "e-mail" or, often, "email") is a store and forward method of composing, sending, storing, and receiving messages over electronic communication systems. The term "e-mail" (as a noun or verb) applies both to the Internet e-mail system based on the Simple Mail Transfer Protocol (SMTP) and to intranet systems allowing users within one organization to e-mail each other. Often these workgroup collaboration organizations may use the Internet protocols for internal e-mail service.

Origins Of E-mail

E-mail predates the inception of the internet, and was in fact a crucial tool in creating the Internet. MIT first demonstrated the Compatible Time-Sharing System (CTSS) in 1961. It allowed multiple users to log into the IBM 7094 from remote dial-up terminals, and to store files online on disk. This new ability encouraged users to share information in new ways. E-mail started in 1965 as a way for multiple users of a time-sharing mainframe computer to communicate. Although the exact history is murky, among the first systems to have such a facility were SDC's Q32 and MIT's CTSS.

E-mail was quickly extended to become network e-mail, allowing users to pass messages between different computers. The messages could be transferred between users on different computers by at least 1966 (it is possible the SAGE system had something similar some time before).

The ARPANET computer network made a large contribution to the development of e-mail. There is one report which indicates experimental inter-system e-mail transfers on it shortly after its creation, in 1969. Ray Tomlinson initiated the use of the @ sign to separate the names of the user and their machine in 1971. The ARPANET significantly increased the popularity of e-mail, and it became the killer app of the ARPANET.

E-mail In Modern Society

Flaming

Many observers bemoan the rise of flaming in written communications. Flaming occurs when one user, usually upset at another user, e-mails the second user an angry and/or antagonistic message. Flaming is assumed to be more common today because of the ease and impersonality of e-mail communications: confrontations in person or over the phone will often be personal-enough to encourage conversants to "hold their tongue," and typing an unhappy message to another is far easier than seeking that other out and confronting him/her directly.

Text/HTML

Both plain text and HTML are used to convey e-mail. While text is certain to be read by all users without problems, many users feel that HTML-based e-mail has a higher aesthetic value. Advantages of HTML include the ability to include inline links and images, set apart previous messages in block quotes, wrap naturally on any display, use emphasis such as underlines and italics, and change font styles. HTML e-mails often include an automatically-generated plain text copy as well, for compatibility reasons.

E-mail Spam

E-mail spam is a subset of spam that involves sending nearly identical messages to numerous recipients by e-mail.

Most definitions of spam are based on the e-mail being Unsolicited Bulk E-mail (UBE). That is, spam is e-mail that is both unsolicited by the recipients and there are many substantively similar e-mails being sent. Spam is usually also unwanted, commercial and sent by automated means and some definitions include those aspects.

Internet Fax

Internet fax uses the internet to receive and send faxes.

Traditional faxing involves sending a scanned copy of a document (a facsimile) from one fax machine to another, over the phone network. Internet faxing (or "online faxing") is a general term which can refer to one of several methods of achieving this over the Internet - with a goal of both reduced costs and increased functionality over traditional faxing.

Depending on the specific method/implementation (see below), advantages of using the internet can include

1. no extra telephone line required for the fax
2. paperless communication, integrated with email
3. send and receive multiple faxes simultaneously
4. reduction in phone costs

While the needs of computer-to-fax communications are well covered, the simplicity of quickly faxing a handwritten document combined with the advantages of email are not.

"iFax" (T.37) was designed for fax machines to directly communicate via email. Faxes are sent as e-mail attachments in a TIFF or PDF format.

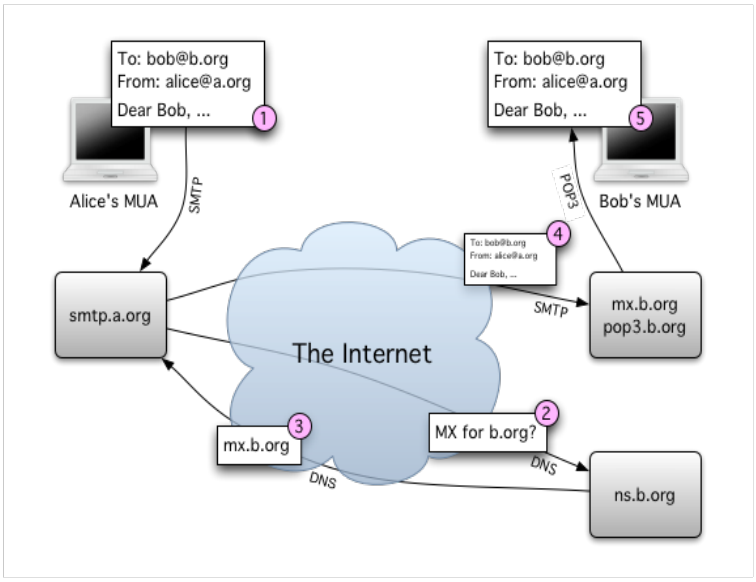
- Fax machine → email message (over Internet) → computer email account
- Fax machine → email message (over Internet) → Fax machine (using email address)

A new fax machine (supporting iFax/T.37) is required, as well as a known email address for the sending and receiving machines. This has limited the standard's use, though a system for looking up a fax's email address based on its phone number is under development [1].

To work with existing fax machines, all iFax machines support standard faxing (requiring a regular phone line). Alternatively, an iFax can be used in conjunction with a fax gateway.

- Fax machine → email message (over Internet) → Fax gateway → Phone line → Fax machine (or vice versa)

How E-mail Works



The diagram above shows a typical sequence of events that takes place when Alice composes a message using her mail user agent (MUA). She types in, or selects from an address book, the e-mail address of her correspondent. She hits the "send" button.

1. Her MUA formats the message in Internet e-mail format and uses the Simple Mail Transfer Protocol (SMTP) to send the message to the local mail transfer agent (MTA), in this case smtp.a.org, run by Alice's Internet Service Provider (ISP).
2. The MTA looks at the destination address provided in the SMTP protocol (not from the message header), in this case bob@b.org. An Internet e-mail address is a string of the form localpart@domain.example, which is known as a Fully Qualified Domain Address (FQDA). The part before the @ sign is the local part of the address, often the username of the recipient, and the part after the @ sign is a domain name. The MTA looks up this domain name in the Domain Name System to find the mail exchange servers accepting messages for that domain.
3. The DNS server for the b.org domain, ns.b.org, responds with an MX record listing the mail exchange servers for that domain, in this case mx.b.org, a server run by Bob's ISP.
4. smtp.a.org sends the message to mx.b.org using SMTP, which delivers it to the mailbox of the user bob.
5. Bob presses the "get mail" button in his MUA, which picks up the message using the Post Office Protocol (POP3).

This sequence of events applies to the majority of e-mail users. However, there are many alternative possibilities and complications to the e-mail system:

- Alice or Bob may use a client connected to a corporate e-mail system, such as IBM's Lotus Notes or Microsoft's Exchange. These systems often have their own internal e-mail format and their clients typically communicate with the e-mail server using a vendor-specific, proprietary protocol. The server sends or receives e-mail via the Internet through the product's Internet mail gateway which also does any necessary reformatting. If Alice and Bob work for the same company, the entire transaction may happen completely within a single corporate e-mail system.

- Alice may not have a MUA on her computer but instead may connect to a webmail service.
- Alice's computer may run its own MTA, so avoiding the transfer at step 1.
- Bob may pick up his e-mail in many ways, for example using the Internet Message Access Protocol, by logging into mx.b.org and reading it directly, or by using a webmail service.
- Domains usually have several mail exchange servers so that they can continue to accept mail when the main mail exchange server is not available.
- E-mails are not secure if e-mail encryption is not used correctly.

It used to be the case that many MTAs would accept messages for any recipient on the Internet and do their best to deliver them. Such MTAs are called open mail relays. This was important in the early days of the Internet when network connections were unreliable. If an MTA couldn't reach the destination, it could at least deliver it to a relay that was closer to the destination. The relay would have a better chance of delivering the message at a later time. However, this mechanism proved to be exploitable by people sending unsolicited bulk e-mail and as a consequence very few modern MTAs are open mail relays, and many MTAs will not accept messages from open mail relays because such messages are very likely to be spam.

Note that the people, e-mail addresses and domain names in this explanation are fictional!

Internet E-Mail Format

The format of Internet e-mail messages is defined in RFC 2822 and a series of RFCs, RFC 2045 through RFC 2049, collectively called Multipurpose Internet Mail Extensions (MIME). Although as of July 13, 2005 (see [3]) RFC 2822 is technically a proposed IETF standard and the MIME RFCs are draft IETF standards, these documents are the de facto standards for the format of Internet e-mail. Prior to the introduction of RFC 2822 in 2001 the format described by RFC 822 was the de facto standard for Internet e-mail for nearly two decades; it is still the official IETF standard. The IETF reserved the numbers 2821 and 2822 for the updated versions of RFC 821 (SMTP) and RFC 822, honoring the extreme importance of these two RFCs. RFC 822 was published in 1982 and based on the earlier RFC 733.

Internet e-mail messages consist of two major sections:

- Header — Structured into fields such as summary, sender, receiver, and other information about the e-mail
- Body — The message itself as unstructured text; sometimes containing a signature block at the end

The header is separated from the body by a blank line.

Internet E-Mail Header

The message header consists of fields, usually including at least the following:

1. From: The e-mail address, and optionally name, of the sender of the message
2. To: The e-mail address[es], and optionally name[s], of the receiver[s] of the message
3. Subject: A brief summary of the contents of the message
4. Date: The local time and date when the message was originally sent

Each header field has a name and a value. RFC 2822 specifies the precise syntax. Informally, the field name starts in the first character of a line, followed by a ":", followed by the value which is continued on non-null subsequent lines that have a space or tab as their first character. Field names and values are restricted to 7-bit ASCII characters. Non-ASCII values may be represented using MIME encoded words.

Note that the "To" field in the header is not necessarily related to the addresses to which the message is delivered. The actual delivery list is supplied in the SMTP protocol, not extracted from the header content. The "To" field is similar to the greeting at the top of a conventional letter which is delivered according to the address on the outer envelope. Also note that the "From" field does not have to be the real sender of the e-mail message. It is very easy to fake the "From" field and let a message seem to be from any mail address. It is possible to digitally sign e-mail, which is much harder to fake. Some Internet service providers do not relay e-mail claiming to come from a domain not hosted by them, but very few (if any) check to make sure that the person or even e-mail address named in the "From" field is the one associated with the connection. Some internet service providers apply e-mail authentication systems to e-mail being sent through their MTA to allow other MTAs to detect forged spam that might apparently appear to be from them.

Other common header fields include:

1. Cc: carbon copy
2. Bcc: Blind Carbon Copy
3. Received: Tracking information generated by mail servers that have previously handled a message
4. Content-Type: Information about how the message has to be displayed, usually a MIME type

Many e-mail clients present "Bcc" (Blind carbon copy, recipients not visible in the "To" field) as a header field. Different protocols are used to deal with the "Bcc" field; at time the entire field is removed, and at times the field remains but the addresses therein are removed. Addresses added as "Bcc" are only added to the SMTP delivery list, and do not get included in the message data. There are differing impressions on the RFC 2822 Protocol pertaining to this subject.

IANA maintains a list of standard header fields.

E-Mail Content Encoding

E-mail was originally designed for 7-bit ASCII. Much e-mail software is 8-bit clean but must assume it will be communicating with 7-bit servers and mail readers. The MIME standard introduced charset specifiers and two content transfer encodings to encode 8 bit data for transmission: quoted printable for mostly 7 bit content with a few characters outside that range and base64 for arbitrary binary data. The 8BITMIME extension was introduced to allow transmission of mail without the need for these encodings but many mail transport agents still don't support it fully. For international character sets, Unicode is growing in popularity.

Saved Message Filename Extension

Most, but not all, e-mail clients save individual messages as separate files, or allow users to do so. Different applications save e-mail files with different filename extensions.

- .eml - This is the default e-mail extension for Mozilla Thunderbird and is used by Microsoft Outlook Express.
- .emlx - Used by Apple Mail.

- .msg - Used by Microsoft Office Outlook.

Messages & Mailboxes

Messages are exchanged between hosts using the Simple Mail Transfer Protocol with software programs called mail transport agents. Users can download their messages from servers with standard protocols such as the POP or IMAP protocols, or, as is more likely in a large corporate environment, with a proprietary protocol specific to Lotus Notes or Microsoft Exchange Servers.

Mail can be stored either on the client, on the server side, or in both places. Standard formats for mailboxes include Maildir and mbox. Several prominent e-mail clients use their own proprietary format and require conversion software to transfer e-mail between them.

When a message cannot be delivered, the recipient MTA must send a bounce message back to the sender, indicating the problem.

Spamming & E-mail Worms

The usefulness of e-mail is being threatened by three phenomena: spamming, phishing and e-mail worms.

Spamming is unsolicited commercial e-mail. Because of the very low cost of sending e-mail, spammers can send hundreds of millions of e-mail messages each day over an inexpensive Internet connection. Hundreds of active spammers sending this volume of mail results in information overload for many computer users who receive tens or even hundreds of junk messages each day.

E-mail worms use e-mail as a way of replicating themselves into vulnerable computers. Although the first e-mail worm affected UNIX computers, the problem is most common today on the more popular Microsoft Windows operating system.

The combination of spam and worm programs results in users receiving a constant drizzle of junk e-mail, which reduces the usefulness of e-mail as a practical tool.

A number of anti-spam techniques mitigate the impact of spam. In the United States, U.S. Congress has also passed a law, the Can Spam Act of 2003, attempting to regulate such e-mail. Australia also has very strict spam laws restricting the sending of spam from an Australian ISP (http://www.aph.gov.au/library/pubs/bd/2003-04/04bd045.pdf), but its impact has been minimal since most spam comes from regimes that seem reluctant to regulate the sending of spam.

Privacy Problems Regarding E-mail

E-mail privacy, without some security precautions, can be compromised because

- e-mail messages are generally not encrypted;
- e-mail messages have to go through intermediate computers before reaching their destination, meaning it is relatively easy for others to intercept and read messages;
- many Internet Service Providers (ISP) store copies of your e-mail messages on their mail servers before they are delivered. The backups of these can remain up to several months on their server, even if you delete them in your mailbox;

- the Received: headers and other information in the e-mail can often identify the sender, preventing anonymous communication.

There are cryptography applications that can serve as a remedy to one or more of the above. For example, Virtual Private Networks or the Tor anonymity network can be used to encrypt traffic from the user machine to a safer network while GPG, PGP or S/MIME can be used for end-to-end message encryption, and SMTP STARTTLS or SMTP over Transport Layer Security/Secure Sockets Layer can be used to encrypt communications for a single mail hop between the SMTP client and the SMTP server.

Another risk is that e-mail passwords might be intercepted during sign-in. One may use encrypted authentication schemes such as SASL to help prevent this.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

The Internet – Instant Messaging

What Is It?

Instant messaging or IM is a form of real-time communication between two or more people based on typed text. The text is conveyed via computers connected over a network such as the Internet.

Overview

Instant messaging requires the use of a client program that hooks up an instant messaging service and differs from e-mail in that conversations are then able to happen in realtime. Most services offer a presence information feature, indicating whether people on one's list of contacts are currently online and available to chat. This may be called a contact list. In early instant messaging programs, each letter appeared as it was typed, and when letters were deleted to correct typos this was also seen in real time. This made it more like a telephone conversation than exchanging letters. In modern instant messaging programs, the other party in the conversation generally only sees each line of text right after a new line is started. Most instant messaging applications also include the ability to set a status message, roughly analogous to the message on a telephone answering machine.

Popular instant messaging services on the public Internet include .NET Messenger Service, AOL Instant Messenger, Excite/Pal, Gadu-Gadu, Google Talk, iChat, ICQ, Jabber, Qnext, QQ, Meetro, Skype, Trillian, Yahoo! Messenger and Rediff Bol Instant Messenger. These services owe many ideas to an older (and still popular) online chat medium known as Internet Relay Chat (IRC).

One can also connect to a instant messaging service with a multiprotocol instant messaging application. (Which allows one instant messenger (IM) client to connect to multiple IM networks.)

Benefits

Instant messaging offers real-time communication and allows easy collaboration, which might be considered more akin to genuine conversation than email's "letter" format. In contrast to e-mail, the parties know whether the peer is available. Most systems allow the user to set an online status or away message so peers are notified when the user is available, busy, or away from the computer. On the other hand, people are not forced to reply immediately to incoming messages. This way, communication via instant messaging can appear to some to be less intrusive than communication via phone. However, not all popular systems allow the sending of messages to people not currently logged on (offline messages), thus removing much of the difference between IM and email.

It is possible to save a conversation, so as to refer to it later. Also, the fact that instant messages typically are logged in a local message history closes the gap to the persistent nature of e-mails and facilitates quick exchange of information such as URLs or document snippets (which can be unwieldy when communicated via telephone).

History

Instant messaging applications began to appear in the 1970s on multi-user operating systems such as UNIX, initially to facilitate communication with other users logged in to the same machine, then on the local network, and subsequently across the internet. Some of these used a peer-to-peer protocol (eg talk, ntalk and ytalk), while others required peers to connect to a server (see talkers and IRC). Because all of these protocols were based inside a console window, most of those discovering the internet in the mid-1990s and equating it with the web tended not to encounter them.

In the last half of the 1980s and into the early 1990s, the Quantum Link online service for Commodore 64 computers offered user-to-user messages between currently connected customers which they called "On-Line Messages" (or OLM for short). Quantum Link's better known later incarnation, America Online, offers a similar product under the name "AOL Instant Messages" (AIM). While the Quantum Link service ran on a Commodore 64, using only the Commodore's PETSCII text-graphics, the screen was visually divided up into sections and OLMs would appear as a yellow bar saying "Message From:" and the name of the sender along with the message across the top of whatever the user was already doing, and presented a list of options for responding. As such, it could be considered a sort of GUI, albeit much more primitive than the later Unix, Windows and Macintosh based GUI IM programs. OLMs were what Q-Link called "Plus Services" meaning they charged an extra per-minute fee on top of the monthly Q-Link access costs.

Modern GUI-based messaging clients began to take off in the late 1990s with ICQ (1996) and AOL Instant Messenger (AIM, 1997) . AOL later acquired Mirabilis, the creators of ICQ. A few years later AOL was awarded two patents for instant messaging by the U.S. patent office. Meanwhile, other companies developed their own applications (Yahoo, MSN, Excite, Ubique, IBM), each with its own proprietary protocol and client; users therefore had to run multiple client applications if they wished to use more than one of these networks.

In 2000, an open source application and open standards-based protocol called Jabber was launched. Jabber servers could act as gateways to other IM protocols, reducing the need to run multiple clients. Modern multi-protocol clients such as Gaim, Trillian, Adium and Miranda can use any of the popular IM protocols without the need for a server gateway.

Recently, many instant messaging services have begun to offer video conferencing features, Voice Over IP (VoIP) and web conferencing services. Web conferencing services integrate both video conferencing and instant messaging capabilities. Some newer instant messaging companies are offering desktop sharing, IP radio, and IPTV to the voice and video features.

The term "instant messenger" is a service mark of Time Warner and may not be used in software not affiliated with AOL in the United States. For this reason, the instant messaging client formerly known as GAIM or gAIM is now only to be referred to as Gaim or gaim.

Cooperation

There have been several attempts to create a unified standard for instant messaging: IETF's SIP (Session Initiation Protocol) and SIMPLE (SIP for Instant Messaging and Presence Leveraging Extensions), APEX (Application Exchange), Prim (Presence and Instant Messaging Protocol), the open XML-based XMPP (Extensible Messaging and Presence Protocol), more commonly known as Jabber and OMA's (Open Mobile Alliance) IMPS (Instant Messaging and Presence Service) created specifically for mobile devices.

Most attempts at creating a unified standard for the major IM providers (AOL, Yahoo! and Microsoft) have failed and each continues to use its own proprietary protocol.

However, while discussions at IETF were stalled, Reuters head of collaboration services, David Gurle (the founder of Microsoft's Real Time Communication and Collaboration business), surprised everybody by signing the first inter-service provider connectivity agreement on September 2003. This historic agreement enabled AIM, ICQ and MSN Messenger users to talk with Reuters Messaging counterparts and vice-versa against an access fee. Following this breakthrough agreement between networks Microsoft, Yahoo! and AOL came to a deal where Microsoft's Live Communication Server 2005 (which is interestingly also used by Reuters for its Reuters Messaging service) users would also have the possibility to talk to public instant messaging users. This deal settled once for all the protocol for interconnectivity in the market as SIP/SIMPLE and established a connectivity fee for accessing public instant messaging clouds. Separately, on October 13, 2005 Microsoft and Yahoo! announced that by (the Northern Hemisphere) summer of 2006 they would interoperate using SIP/SIMPLE which is followed on December 2005 by the AOL and Google strategic partnership deal where Google Talk users would be able to talk with AIM and ICQ users provided they have an identity at AOL.

There are two ways to combine the many disparate protocols:

1. One way is to combine the many disparate protocols inside the IM client application. Examples include iChat, Trillian, Gaim, Fire, Proteus, Miranda IM, Adium, Everybuddy, Ayttm, Kopete, Centericq, BitlBee, Windows Messenger, and IMVITE.
2. The other way is to combine the many disparate protocols inside the IM server application. This approach moves the task of communicating to the other services to the server. Clients need not know or care about other IM protocols. For example, LCS 2005 Public IM Connectivity. This approach is popular in Jabber/XMPP servers however the so-called transport projects suffer the same reverse engineering difficulties as any other project involved with closed protocols or formats.

Some approaches, such as that adopted by the Sonork enterprise IM software or the Jabber/XMPP network or Winpopup LAN Messenger, allow organizations to create their own private instant messaging network by enabling them to limit access to the server (often with the IM network entirely behind their firewall) and administer user permissions. Other corporate messaging systems, like the Medianet Innovations MIC, allow registered users to also connect from outside the corporation LAN, by using a secure firewall-friendly HTTPS based protocol. Typically, a dedicated corporate IM server has several advantages such as pre-populated contact lists, integrated authentication, and better security and privacy.

Some networks have made changes to prevent them from being utilized by such multi-network IM clients. For example, Trillian had to release several revisions and patches to allow its users to access the MSN, AOL, and Yahoo! networks, after changes were made to these networks. The major IM providers typically cite the need for formal agreements as well as security concerns as reasons for making these changes.

Mobile Instant Messaging

Mobile Instant Messaging abbreviated as MIM is a presence enabled messaging service that aims to transpose the desktop messaging experience to the usage scenario of being on the move. While several of the core ideas of the desktop experience on one hand apply to a connected mobile device, others do not: Users usually only look at their phone's screen -- presence status changes might occur under different circumstances as happens at the desktop, and several functional limits exist based on the fact that the vast majority of mobile communication devices are chosen by their users to fit into the palm of their hand.

Some of the form factor and mobility related differences need to be taken into account in order to create a really adequate, powerful and yet convenient mobile experience: radio bandwidth, memory size, availability of media formats, keypad based input, screen output, CPU performance and battery power are core issues that desktop device users and even nomadic users with connected notebooks are usually not exposed to.

Several formerly untackled issues have been identified and addressed within IMPS. This standard (IMPS) was developed as part of an early mobile telephone industry initiative to kick off a broader usage of mobile instant messaging . The Open Mobile Alliance has taken over this standard, formerly called Wireless Village, as IMPS V1.0 in November 2002. Since then this standards has been further developed to IMPS V1.3, the latest candidate for release, and is expected to be released before the end of 2006.

There are downloadable mobile applications offered by different independent developers that allow users to chat within public (MSN, Yahoo! , Google Talk, AIM, ICQ) and corporate (LCS, Sametime, Reuters) IM services from mobile devices. The example of such multi-IM software is IM+ Instant Messenger available for BlackBerry, Symbian OS, Windows Mobile, BREW, J2ME, i-mode and Garnet OS (former Palm OS).

Among the advantages of using such IM clients over SMS are: IM clients use data instead of SMS text messages; IM-like chat mode, faster and quicker messaging. Some IM software allows group communication.

Several large scale mobile telephone industry companies are planning to jointly deliver a ubiquitous, interoperable presence enabled messaging service, built according to interoperability recommendations developed in the GSM Association. Considering these organisations are jointly representing approximately 1.5 billion active Short Text Messaging (SMS) users, it remains to be seen if such an initiative may also help to drive the different industry factions to agree on a truly interoperable approach at least for Mobile Instant Messaging sometime in the not too far future.

In the meantime, other developments have proposed usage of downloadable applications with the intention to create their own approach to IM that runs on most mobile phones worldwide. Essentially, several of these clients are Java applications, such as MXit or Thumb Messenger, that are instantly downloaded and then connected to back-end servers through GPRS/3G Internet Channels. Some of the implementations can connect to other IM services like Jabber, Google Talk, MSN Messenger and AOL's AIM, Rediff Bol Instant Messenger and ICQ.

Effects On People With Hearing Loss

Instant messaging opens new methods of spontaneous communication for people with hearing loss. It is considered by many a powerful way to allow equal opportunities in communication, without the aid of special devices or services design for users with hearing loss.

Friend-To-Friend Networks

Instant Messaging may be done in a Friend-to-friend network, in which each node connects to the friends on the friendslist. This allows to communicate to friends of friends and build chatrooms for instant messages with all friends on that network.

User Base

Note that the numbers listed in this section might be very speculative. This is caused by the fact that some numbers are given by the owners of the instant mesaging system or commercial vendors of a part of a distributed system. These companies want to enlarge this number in order to receive more advertisement earnings or to acquire more customers. Another reason why these numbers might be speculative is because some instant messaging systems are distributed amongst different instances and thus impossible to easily measure (e.g. Jabber).

- AIM: 53 million active users (September 2006), "over 100 million" total (January 2006).
- Jabber: between 40 and 50 million (January 2007). Note that this number is based on calculations of Jabber Inc ("nearly 10 million open source users") which differ from those of Process-One ("Our total deployments account for more than 20 millions of accounts"). Process-One is a company providing services based on the Jabber server software ejabberd. Accordingly, as there are many other open source servers (some also with companies behind it), the number provided by Jabber Inc is probably too small. If we presume ejabberd has a 40% market share amongst public and non-public open source server deployments, there are 50 million of users using an open source server. This would mean, including Jabber Inc's numbers, that there are around 90 million of Jabber users instead of 50 million.
- Ebuddy: 35 million users (including 4 million mobile) (October 2006)
- Windows Live Messenger: 27.2 million active (September 2006), 155 million total (April 2005).
- Yahoo! Messenger: 22 million users (September 2006).
- QQ: 20 million peak online users, 221 million "active"[4] (July 2006).
- Sametime: 15 million (enterprise) users (undated)
- Skype: 9 million peak online (January 2007), 137 million total (January 2007).
- Xfire: 6.1 million users (January 2007)
- Gadu-Gadu: 5.6 million users (June 2006).
- ICQ: 4 million active (September 2006).
- Paltalk: 3.3 million unique visitors per month (August 2006).
- MXit: 3 million users (majority in South Africa, more than 200,000 international) (31 January 2007). Note that these users are part of the Jabber user base as MXit federates with the Jabber network.
- PSYC: 1 million users, daily (majority in Brazil) (February 2007). Total amount of users cannot be estimated due to the decentralized nature of the protocol.
- Meebo: 1 million users (October 2006)

- Tjat : 4.245 users 0.63 million active users
- iGo Incognito: unknown; no objective numbers available. Based in Auckland, New Zealand
- Orooni: unknown; no objective numbers available. Based in Oslo, Norway

The Internet – Weblogs & Blogs

What Is A Weblog?

Weblog software (also called blog software or blogware) is a category of software which consists of a specialized form of Content Management Systems specifically designed for creating and maintaining weblogs.

Server Models

Many weblog applications are available for users to download and install on their own systems. A wide variety of licenses are used by user hosted weblog software. Some of these are open-source software that can be used, modified, and redistributed freely, with no usage restrictions. Others are proprietary software that may be licensed for a fee or have versions available free of charge.

Other weblog applications are offered only through their developers' hosts, either free of charge or for a fee. These typically include hosting service for the published blog itself, but some offer the option of using this hosted software to update a blog published

Clients

Maintenance through the Internet is a nearly universal feature of weblog software. This is usually done through a browser-based interface, enabling authors to create and update content on the site. Most software supports the use of external client software to update content using common APIs such as the MetaWeblog API and the Atom Publishing Protocol.

Features

- Title, the main title, or headline, of the post.
- Body, main content of the post.
- Permalink, the URL of the full, individual article.
- Post Date, date and time the post was published.

A blog entry optionally includes the following:

- Comments - Comments are a way to provide discussion on blog entries. Readers can leave a comment on a post, which can correct errors or contain their opinion on the post or the post's subject. Services like coComment aim to ease discussion through comments, by allowing tracking of them.
- Categories (or tags) - subjects that the entry discusses
- Trackback and or pingback - links to other sites that refer to the entry

Other Applications

Most weblogs have features such as facilitating authoring and editing of blog posts or articles, various linking and web syndication features, and the ability to easily publish the blog to the world wide web. Some services or organizations are also creating weblog applications with extended features to aid communication, such as the wiki capabilities in Socialtext and Traction TeamPage.

Many weblog applications allow the user to define static pages of content which can often be placed into a hierarchy or tree. Pages differ from blog posts in that the content is largely static and not time related. Pages are often used to present information about the blog and its authors. Extensive use of pages can result in a blog that looks more like a website.

Most weblog applications support English and many other languages. The user selects a language during installation.

Weblog applications usually offer web syndication service either in the form of RSS or Atom. This allows for other software such as feed aggregators to maintain a current summary of the blog's content.

Post moderation requires the people who want to comment on articles that are posted on a blog to be approved before the comments are visible to the world. It could also mean in some cases where multiple people have accounts and the ability to post new items to the blog that new content must be approved by a moderator or administrator before it shows up on the main page. Weblog applications use various user account systems that allow readers to post comments to a particular blog. For instance, users with Blogger accounts may comment on any Blogger blog. Other weblog applications allow users to post content or comments only to blogs where they have an account.

The Post API can vary greatly depending on the system that you are using. Some types of blogware have plugins for Firefox that integrate into the browser's menus so that right-clicking on selected text on any given webpage will bring up a small window that allows the user to post to their blog. Other types of blogware that do not have this type of interface require a person to fill out a form online. The form that is required for posting material to a blog depends on the type of blogware. Some types such as Movable Type contain a greater number of form fields and choices than ones such as Blogger.

All types of blogware support adding thumbnail images within blog posts. Photo blogging is a separate genre of blogging that deals primarily with images.

Documentation & Support

Different blogware packages feature varying levels of community support and documentation. Because the installation of some types of software requires an advanced knowledge of computer administration, community support and documentation can be very helpful. The web servers and database software can be more difficult to install than the blogware itself. Also a strong and active community surrounding the blogware gives advice on integrating the blogware into a personal site.

Examples

A partial list of notable weblog software follows:

User-Hosted

Software packages installed by weblog authors to run on their own systems:

Free/Libre/Open-Source Software

These software packages are offered under a Free/Libre/Open-Source Software licence. Therefore they are free for everyone to use. Also, commercial support contracts are often available.

- Apache Roller (Java-based)
- b2evolution (PHP/MySQL)
- blosxom (Perl)
- PyBlosxom (Python)
- Dotclear (PHP/MySQL)
- Drupal (PHP/MySQL)
- Frog CMS (PHP/MySQL)
- Elgg (Linux, Apache, MySQL, and PHP)
- Habari (PHP/MySQL,SQLite)
- Livejournal (Perl) (Also available, developer hosted)
- LifeType (PHP/MySQL)
- Movable Type (also offered in developer-hosted form as TypePad)
- Nucleus CMS (PHP)
- phsBlog (PHP/MySQL)
- Picoplog (PHP) (photoblog)
- Pixelpost (PHP/MySQL) (photoblog)
- Serendipity (PHP/MySQL,PostgreSQL,SQLite)
- SimplePHPBlog (PHP/TXT)
- Slash (Perl/MySQL)
- Subtext (C#/ASP.NET)
- Textpattern (PHP/MySQL)
- Typo (Ruby on Rails)
- WordPress (PHP/MySQL) - also offered in developer-hosted form as Wordpress.com

Proprietary Software

These packages are under a proprietary license. They may require the purchase of a license key to use them. The specific licensing terms vary but some are free for personal or non-commercial use.

- Community Server (also offered in developer-hosted form as CommunityServer.com)
- ExpressionEngine
- Radio UserLand
- Traction TeamPage
- Windows Live Writer (Free of charge)

Unknown License

- Battle Blog
- Blogsphere

Developer-Hosted

Software services operated by the developer, requiring no software installation for the weblog author:

- Battle Blog
- Blog.com
- Blogabond
- Blogger
- Blogging Systems

- Blogonize
- BlogSavy
- Blue Kaffee
- Blurty
- DeadJournal
- GreatestJournal
- LiveJournal
- InsaneJournal
- MySpace
- Open Diary
- Quillpill
- Square Space
- TypePad
- Typo
- Windows Live Spaces
- Wordpress.com (hosted version of WordPress)
- Xanga
- Vox

What Is A Blog?

A blog (an abridgment of the term web log) is a website, usually maintained by an individual, with regular entries of commentary, descriptions of events, or other material such as graphics or video. Entries are commonly displayed in reverse chronological order. "Blog" can also be used as a verb, meaning to maintain or add content to a blog.

Many blogs provide commentary or news on a particular subject; others function as more personal online diaries. A typical blog combines text, images, and links to other blogs, web pages, and other media related to its topic. The ability for readers to leave comments in an interactive format is an important part of many blogs. Most blogs are primarily textual, although some focus on art (artlog), photographs (photoblog), sketchblog, videos (vlog), music (MP3 blog), audio (podcasting) are part of a wider network of social media. Micro-blogging is another type of blogging which consists of blogs with very short posts. As of December 2007, blog search engine Technorati was tracking more than 112 million blogs.

Blog Types

There are many different types of blogs, differing not only in the type of content, but also in the way that content is delivered or written.

- Personal blogs - The personal blog, an on-going diary or commentary by an individual, is the traditional, most common blog. Personal bloggers usually take pride in their blog posts, even if their blog is never read by anyone but them. Blogs often become more than a way to just communicate; they become a way to reflect on life or works of art. Blogging can have a sentimental quality. Few personal blogs rise to fame and the mainstream, but some personal blogs quickly garner an extensive following.

- Corporate Blogs - A blog can be private, as in most cases, or it can be for business purposes. Blogs, either used internally to enhance the communication and culture in a corporation or externally for marketing, branding or PR purposes are called corporate blogs.

- By Media Type
  - A blog comprising videos is called a vlog, one comprising links is called a linklog, a site containing a portfolio of sketches is called a sketchblog or one comprising photos is called a photoblog. Blogs with shorter posts and mixed media types are called tumblelogs.
  - An Artlog is a form of art sharing and publishing in the format of a blog, but differentiated by the predominant use of and focus on Art work rather than text.
  - A rare type of blog hosted on the Gopher Protocol is known as a Phlog

- By Device - Blogs can also be defined by which type of device is used to compose it. A blog written by a mobile device like a mobile phone or PDA could be called a moblog. One early blog was Wearable Wireless Webcam, an online shared diary of a person's personal life combining text, video, and pictures transmitted live from a wearable computer and EyeTap device to a web site. This practice of semi-automated blogging with live video together with text was referred to as sousveillance. Such journals have been used as evidence in legal matters.

- By Genre - Some blogs focus on a particular subject, such as political blogs, travel blogs, fashion blogs, project blogs, education blogs, niche blogs, classical music blogs,quizzing blogs and legal blogs (often referred to as a blawgs) or dreamlogs. While not a legitimate type of blog, one used for the sole purpose of spamming is known as a Splog.

Community & Cataloging

- The Blogosphere - The collective community of all blogs is known as the blogosphere. Since all blogs are on the internet by definition, they may be seen as interconnected and socially networked. Discussions "in the blogosphere" have been used by the media as a gauge of public opinion on various issues. A collection of local blogs is sometimes referred to as a bloghood.

- Blog Search Engines - Several blog search engines are used to search blog contents, such as Bloglines, BlogScope, and Technorati. Technorati, which is among the most popular blog search engines, provides current information on both popular searches and tags used to categorize blog postings. Research community is working on going beyond simple keyword search, by inventing news ways to navigate through huge amounts of information present in the blogosphere, as demonstrated by projects like BlogScope.

- Blogging Communities & Directories - Several online communities exist that connect people to blogs and bloggers to other bloggers, including BlogCatalog and MyBlogLog.

- Blogging & Advertising - It is common for blogs to feature advertisements either for the financial benefit of the blogger, or to promote the blogger's favorite causes. The popularity of blogs has also given rise to "fake blogs" in which a company will create a fictional blog as a marketing tool to promote a product.

Popularity

Recently, researchers have analyzed the dynamics of how blogs become popular. There are essentially two measures of this: popularity through citations, as well as popularity through affiliation (i.e. blogroll). The basic conclusion from studies of the structure of blogs is that while it takes time for a blog to become popular through blogrolls, permalinks can boost popularity more quickly, and are perhaps more indicative of popularity and authority than blogrolls, since they denote that people are actually reading the blog's content and deem it valuable or noteworthy in specific cases.

The blogdex project was launched by researchers in the MIT Media Lab to crawl the Web and gather data from thousands of blogs in order to investigate their social properties. It gathered this information for over 4 years, and autonomously tracked the most contagious information spreading in the blog community, ranking it by recency and popularity. It can thus be considered the first instantiation of a memetracker. The project is no longer active, but a similar function is now served by tailrank.com.

Blogs are given rankings by Technorati based on the number of incoming links and Alexa Internet based on the web hits of Alexa Toolbar users. In August 2006, Technorati found that the most linked-to blog on the internet was that of Chinese actress Xu Jinglei. Chinese media Xinhua reported that this blog received more than 50 million page views, claiming it to be the most popular blog in the world. Technorati rated Boing Boing to be the most-read group-written blog.

Gartner forecasts that blogging will peak in 2007, leveling off when the number of writers who maintain a personal website reaches 100 million. Gartner analysts expect that the novelty value of the medium will wear off as most people who are interested in the phenomenon have checked it out, and new bloggers will offset the number of writers who abandon their creation out of boredom. The firm estimates that there are more than 200 million former bloggers who have ceased posting to their online diaries, creating an exponential rise in the amount of "dotsam" and "netsam" — that is to say, unwanted objects on the Web.

Blurring With The Mass Media

Many bloggers, particularly those engaged in participatory journalism, differentiate themselves from the mainstream media, while others are members of that media working through a different channel. Some institutions see blogging as a means of "getting around the filter" and pushing messages directly to the public. Some critics worry that bloggers respect neither copyright nor the role of the mass media in presenting society with credible news. Bloggers and other contributors to user generated content are behind Time magazine naming their 2006 person of the year as "you".

Many mainstream journalists, meanwhile, write their own blogs — well over 300, according to CyberJournalist.net's J-blog list. The first known use of a weblog on a news site was in August 1998, when Jonathan Dube of The Charlotte Observer published one chronicling Hurricane Bonnie.

Some bloggers have moved over to other media. The following bloggers (and others) have appeared on radio and television: Duncan Black (known widely by his pseudonym, Atrios), Glenn Reynolds (Instapundit), Markos Moulitsas Zúniga (Daily Kos), Alex Steffen (Worldchanging) and Ana Marie Cox (Wonkette). In counter-point, Hugh Hewitt exemplifies a mass media personality who has moved in the other direction, adding to his reach in "old media" by being an influential blogger.

Blogs have also had an influence on minority languages, bringing together scattered speakers and learners; this is particularly so with blogs in Gaelic languages, whose creators can be found as far away from traditional Gaelic areas as Kazakhstan and Alaska. Minority language publishing (which may lack economic feasibility) can find its audience through inexpensive blogging.

There are many examples of bloggers who have published books based on their blogs, e.g., Salam Pax, Ellen Simonetti, Jessica Cutler, ScrappleFace. Blog-based books have been given the name blook. A prize for the best blog-based book was initiated in 2005, the Lulu Blooker Prize. However success has been elusive offline, with many of these books not selling as well as their blogs. Only blogger Tucker Max cracked the New York Times Bestseller List.

Blogging Consequences

The emergence of blogging has brought a range of legal liabilities and other often unforeseen consequences. One area of concern is the issue of bloggers releasing proprietary or confidential information. Another area of concern is blogging and defamation. A third area of concern is employees who write about aspects of their place of employment or their personal lives, and then face loss of employment or other adverse consequences. A number of examples of blogging and its sometimes negative or unforeseen consequences are cited here.

Defamation Or Liability

Several cases have been brought before the national courts against bloggers concerning issues of defamation or liability. The courts have returned with mixed verdicts. Internet Service Providers (ISPs), in general, are immune from liability for information that originates with Third Parties (U.S. Communications Decency Act and the EU Directive 2000/31/EC).

In John Doe v. Patrick Cahill, the Delaware Supreme Court held that stringent standards had to be met to unmask anonymous bloggers, and also took the unusual step of dismissing the libel case itself (as unfounded under American libel law) rather than referring it back to the trial court for reconsideration. In a bizarre twist, the Cahills were able to obtain the identity of John Doe, who turned out to be the person they suspected: the town's mayor, Councilman Cahill's political rival. The Cahills amended their original complaint, and the mayor settled the case rather than going to trial.

In 2004, eight Royal Dutch Shell Group companies collectively obtained an "Interim Injunction and Restraining Order" in Malaysia against a Shell whistleblower and former employee, Dr John Huong, a Malaysian geologist. Dr Huong had allegedly posted defamatory material on a weblog hosted in North America, royaldutchshellplc.com. The weblog site was owned and operated by British national and long-term Shell critic, Alfred Donovan. Additional legal actions were initiated against Dr Huong in 2006 in response to publications on Donovan's weblog sites in 2005 and 2006. Those actions included a "Notice to Show Cause" relating to a "contempt of court," which was potentially punishable by imprisonment.

In January 2007, two prominent Malaysian political bloggers, Jeff Ooi and Ahiruddin Attan were sued by pro-government newspaper, The New Straits Times Press (Malaysia) Berhad, Kalimullah bin Masheerul Hassan, Hishamuddin bin Aun and Brenden John a/l John Pereira over an alleged defamation. The plaintiff was supported by the Malaysian government. Following the suit, the Malaysian government proposed to "register" all bloggers in Malaysia in order to better control parties against their interest. This is the first such legal case against bloggers in the country.

In Britain, a college lecturer contributed to a blog in which she referred to a politician (who

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

had also expressed his views in the same blog) using various uncomplimentary names, including referring to him as a "Nazi". The politician found out the real name of the lecturer (she wrote under a pseudonym) via the ISP and successfully sued her for £10,000 in damages and £7,200 costs.

In the United States blogger Aaron Wall was sued by Traffic Power for defamation and publication of trade secrets in 2005. According to Wired Magazine, Traffic Power had been "banned from Google for allegedly rigging search engine results." Wall and other "white hat" search engine optimization consultants had exposed Traffic Power in what they claim was an effort to protect the public. The case was watched by many bloggers because it addressed the murky legal question of who's liable for comments posted on blogs.

Employment

In general, attempts at hiding the blogger's name and/or the place of employment in anonymity have proved ineffective at protecting the blogger. Employees who blog about elements of their place of employment raise the issue of employee branding, since their activities can begin to affect the brand recognition of their employer.
In fall 2004, Ellen Simonetti was fired for what was deemed by her employer to be inappropriate material on her blog. She subsequently wrote a book based on her blog.
In fall 2004, Ellen Simonetti was fired for what was deemed by her employer to be inappropriate material on her blog. She subsequently wrote a book based on her blog.

Ellen Simonetti, a Delta Air Lines flight attendant, was fired by the airline for photos of herself in uniform on an airplane and comments posted on her blog "Queen of Sky: Diary of a Flight Attendant" which her employer deemed inappropriate. This case highlighted the issue of personal blogging and freedom of expression vs. employer rights and responsibilities, and so it received wide media attention. Simonetti took legal action against the airline for "wrongful termination, defamation of character and lost future wages". The suit is postponed while Delta is in bankruptcy proceedings (court docket).

In the spring of 2006, Erik Ringmar, a tenured senior lecturer at the London School of Economics was ordered by the convenor of his department to "take down and destroy" his blog in which he discussed the quality of education at the school.

Mark Cuban, owner of the Dallas Mavericks, was recently fined during the 2006 NBA playoffs for criticizing NBA officials on the court and in his blog.

Mark Jen was terminated in 2005 after a mere 10 days of employment at Google for discussing corporate secrets on his personal blog.

In India, blogger Gaurav Sabnis resigned from IBM after his posts exposing the false claims of a management school, IIPM, led to management of IIPM threatening to burn their IBM laptops as a sign of protest against him.

Jessica Cutler, aka "The Washingtonienne", blogged about her sex life while employed as a congressional assistant. After the blog was discovered and she was fired, she wrote a novel based on her experiences and blog: The Washingtonienne: A Novel. Cutler is presently being sued by one of her former lovers in a case that could establish the extent to which bloggers are obligated to protect the privacy of their real life associates.

Catherine Sanderson, aka Petite Anglaise, lost her job in Paris at a British accountancy firm as a consequence of blogging. Although given in the blog in a fairly anonymous manner, some of the descriptions of the firm and some of its people were less than flattering. Sanderson later won a compensation claim case against the British firm, however.

On the other hand, Penelope Trunk, writing in the Globe in 2006, was one of the first to point out that a large portion of bloggers are professionals, and a well written blog can actually help attract employers.

Political Dangers

Blogging can sometimes have unforeseen consequences in politically sensitive areas. Blogs are much harder to control than broadcast or even print media. As a result totalitarian and authoritarian regimes often seek to suppress blogs, or to punish those who maintain them.

In Singapore, two ethnic Chinese were imprisoned under the country’s anti-sedition law for posting anti-Muslim remarks in their weblogs.

Egyptian blogger Kareem Amer was charged with insulting the Egyptian president Hosni Mubarak and an Islamic institution through his online blog. It is the first time in the history of Egypt that a blogger was prosecuted. After a brief trial session that took place in Alexandria, the blogger was found guilty and sentenced to prison terms of three years for insulting Islam and inciting sedition, and one year for insulting Mubarak.

Egyptian blogger Abdel Monem Mahmoud was arrested in April 2007 for anti-government writings in his blog. Monem is a member of the banned Muslim Brotherhood.

After expressing opinions in his personal weblog about the state of the Sudanese armed forces, Jan Pronk, United Nations Special Representative for the Sudan, was given three days notice to leave Sudan. The Sudanese army had demanded his deportation.

Personal Safety

One unfortunate consequence of blogging is the possibility of attacks or threats against the blogger, sometimes without apparent reason. Kathy Sierra, author of the innocuous blog Creating Passionate Users, was the target of such vicious threats and misogynistic insults that she canceled her keynote speech at a technology conference in San Diego, fearing for her safety. While a blogger's anonymity is often tenuous, internet trolls who would attack a blogger with threats or insults can be emboldened by anonymity. Sierra and supporters initiated an online discussion aimed at countering abusive online behavior and developed a blogger's code of conduct.

History

The term "weblog" was coined by Jorn Barger on 17 December 1997. The short form, "blog," was coined by Peter Merholz, who jokingly broke the word weblog into the phrase we blog in the sidebar of his blog Peterme.com in April or May of 1999. This was quickly adopted as both a noun and verb ("to blog," meaning "to edit one's weblog or to post to one's weblog").

Origins

Before blogging became popular, digital communities took many forms, including Usenet, commercial online services such as GEnie, BiX and the early CompuServe, e-mail lists and Bulletin Board Systems (BBS). In the 1990s, Internet forum software, such as WebEx,

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

created running conversations with "threads". Threads are topical connections between messages on a metaphorical "corkboard".

The modern blog evolved from the online diary, where people would keep a running account of their personal lives. Most such writers called themselves diarists, journalists, or journalers. Justin Hall, who began eleven years of personal blogging in 1994 while a student at Swarthmore College, is generally recognized as one of the earliest bloggers, as is Jerry Pournelle. Dave Winer's Scripting News is also credited with being one of the oldest and longest running weblogs.

Early weblogs were simply manually updated components of common websites. However, the evolution of tools to facilitate the production and maintenance of web articles posted in reverse chronological order made the publishing process feasible to a much larger, less technical, population. Ultimately, this resulted in the distinct class of online publishing that produces blogs we recognize today. For instance, the use of some sort of browser-based software is now a typical aspect of "blogging". Blogs can be hosted by dedicated blog hosting services, or they can be run using blog software, such as WordPress, Movable Type, Blogger or LiveJournal, or on regular web hosting services.

Rise In Popularity

After a slow start, blogging rapidly gained in popularity. Blog usage spread during 1999 and the years following, being further popularized by the near-simultaneous arrival of the first hosted blog tools:

- Open Diary launched in October 1998, soon growing to thousands of online diaries. Open Diary innovated the reader comment, becoming the first blog community where readers could add comments to other writers' blog entries.
- Brad Fitzpatrick, a well known blogger started LiveJournal in March 1999.
- Andrew Smales created Pitas.com in July 1999 as an easier alternative to maintaining a "news page" on a website, followed by Diaryland in September 1999, focusing more on a personal diary community.
- Evan Williams and Meg Hourihan (Pyra Labs) launched blogger.com in August 1999 (purchased by Google in February 2003)

Blogging Becomes Mainstream

By 2004, the role of blogs became increasingly mainstream, as political consultants, news services and candidates began using them as tools for outreach and opinion forming. Blogging was established by politicians and political candidates to express opinions on war and other issues and cemented blogs' role as a news source. Even politicians not actively campaigning, such as the UK's Labour Party's MP Tom Watson, began to blog to bond with constituents.

In January 2005, Fortune magazine listed eight bloggers that business people "could not ignore": Peter Rojas, Xeni Jardin, Ben Trott, Mena Trott, Jonathan Schwartz, Jason Goldman, Robert Scoble, and Jason Calacanis.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited.**io

The Internet – Podcasting

High-Level

A podcast is a series of digital-media files which are distributed over the Internet using syndication feeds for playback on portable media players and computers. The term podcast, like broadcast, can refer either to the series of content itself or to the method by which it is syndicated; the latter is also called podcasting. The host or author of a podcast is often called a podcaster.

The term is a portmanteau of the words "iPod" and "broadcast", the Apple iPod being the brand of portable media player for which the first podcasting scripts were developed. These scripts allowed podcasts to be automatically transferred to a mobile device after they are downloaded.

Though podcasters' web sites may also offer direct download or streaming of their content, a podcast is distinguished from other digital media formats by its ability to be syndicated, subscribed to, and downloaded automatically when new content is added, using an aggregator or feed reader capable of reading feed formats such as RSS or Atom.

Variants

- Autocasting (the automatic generation of podcasts from text-only sources)
- Blogcasting (the blogging Podcast)
- Mediacasting (any distribution of audio/video media files utilizing RSS)
- Mobilecast (podcasting to mobile phones)
- Vodcasting (video podcasting)
- Narrowcasting (podcasting is a form of narrowcasting)
- Peercasting (peercasting allows live streams to be redistributed by the viewers/listener, greatly reducing bandwidth needs for the originating broadcaster)

The Author of this prelim study would like to suggest an additional new variant, "Textcasting", an eBook reading mechanism people would actually use

Receiving & Using Podcasts

Making full use of podcasts' syndication features requires appropriate software, often referred to as a podcatching client. The feeds are usually distributed using RSS or Atom protocols to the podcatching client. The dominant podcatching client is Apple's iTunes player. However, there are alternatives, including Microsoft's Zune Marketplace, Juice (multiplatform), Doppler (Windows), Podget (Linux) and Podracer (Linux). Some established audio players, such as Amarok, Winamp and Mediamonkey also offer (sometimes limited) podcatching functionality.

Many podcasts also allow users to direct download, by giving a link to the audio file in an RSS feed or web page.

Podcast listeners can listen in one of three ways: through an MP3 player, on a computer using media player software, or with VoIP technology by calling to a virtual phone number. By dialing a phone number, you can hear a menu of available podcasts and features. Find more about it in the external links section.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

Uses Of Podcasts

Public Services

- Unofficial audio tours of museums (musecast)
- Official cultural or historic audio tours of cities
- A way for news organizations to distribute audio or video as an addition to their existing text (or mostly text) news products. For example, Wikinews began to podcast its News Briefs in 2005. Companies are also using podcasts as a way to distribute their multimedia news to journalists and consumers through companies like MultiVu. In 2006, the online magazine Slate began textcasting articles to their readers, by attaching a written article to a blank audio file and delivering the content to readers through their regular podcasting mechanism.
- Advocacy. The 5,500 locked out staff (editors, journalists, technicians, hosts, etc.) of the Canadian Broadcasting Corporation were podcasting news and other programming during August and September of 2005.
- Youth media. Podcasting has become a way for youth media organizations, such as Youth Radio (Youth Radio site), to bring youth perspectives to a wider audience.
- Public libraries can podcast local publications free of Copyright, offering spoken word alternatives to the visually impaired. Non-profit organizations like Assistive Media podcast readings of short-format magazine articles for visually impaired readers.
- Law enforcement. The Chicago Police Department has a free video podcast of its half-hour weekly news magazine called "CrimeWatch," which airs on local TV. It documents community policing (CAPS) success stories.

Education & Academia

Podcasts enable students and teachers to share information with anyone anytime. If a student is absent, he or she can download the podcast of the recorded lesson. Teachers may also create podcasts to be used as a preparation tool for students. This would be pedagogically equivalent to having students read a text before a lesson. It can be a tool for teachers or administrators to communicate curriculum, assignments and other information with parents and the community. Teachers can record book talks, vocabulary or foreign language lessons, international pen pal letters (podcast pals!), music performance, interviews, debates. Podcasting can be a publishing tool for student oral presentations. Video podcasts can be used in all these ways as well.

- Mobile Learning: Podcasting can be categorised as an m-learning strategy for teaching and learning. In 2004 Musselburgh Grammar School pioneered podcast lessons with foreign language audio revision and homework, other pioneers include The Room 208 Podcast, Radio WillowWeb, and Room 613 Talk. In the second half of 2005, a Communication Studies course at the University of Western Australia (iGeneration: Digital Communication and Participatory Culture) used student-created podcasts as the main assessment item. Podcasts have proven beneficial in early elementary education as well. In 2005 Students in the Write was created for second grade students at Morse Elementary School in Tarrytown, NY. By providing students with an authentic audience, teachers noticed significantly increased motivation to write. Students were also found to improve fluency and listening skills. On the 21st February 2006 Lance Anderson, Dr. Chris Smith (the Naked Scientist), Nigel Paice and Debbie McGowan took part in the first podcast forum at Cambridge University. The event was hosted by the Centre for Applied Research in Educational Technologies. See "A Journey to Podcasting" at http://www.lanecc.edu/fpd/grants/sabbatical/paid/documents/AnnickToddreport2.pdf

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

- Journalism Education: School podcasts can be created to expose students to journalism and new-media concepts. Regularly released "news" podcasts can be released by a school group such as the AquiCastpodcast.
- Academic Journal Digests: The Society of Critical Care Medicine has a podcast used to update clinicians with summaries of important articles, as well as interviews.
- Professional Development: Professional development podcasts exist for educators. Some podcasts may be general in nature such as thePodcast for Teachers, or may be slightly more specific such as the Smart Board Podcastwhich focuses on the use of interactive white boards in the classroom.
- Religion: Godcasting has been used by many religious groups. Many churches produce podcasts of talks and sermons. Disciples with Microphones provides podcasts relating to the Catholic Church.
- Tutorials: A tutorial on almost any subject can be created as either an audio podcast or video vodcast. Through screencasting, many video podcasts, such as ScreenCastsOnline, demonstrate how to use software and operating systems.

Entertainment

- Television commentary. Battlestar Galactica writer and executive producer Ronald D. Moore creates commentary podcasts for each new episode of Battlestar Galactica (download audio commentary). Other television shows such as Doctor Who have since followed suit.
- As a platform for fan DVD-style commentary tracks (Audio commentary). Enables fans to add their own comments and thoughts to any of their favourite films.
- Sports. In 2005, unofficial podcasts for major sports teams launched, providing fans both in and outside of the teams' direct broadcast areas with on-demand commentary. Pioneers include Cubscast. The Cubscast founders also formed the first city-specific sports podcast network, hosting one podcast for each major Chicago team at Chicagosportscasts.com.
- Pornography. Porncasting and podnography are sometimes used to refer to pornography in podcasts.

News

- Newspapers. Newspapers use podcasts to broadcast audio content from print interviews and drive traffic to their websites. The San Francisco Chronicle is believed to be the first major daily newspaper to start podcasting using an external website, in Feb 2005. Hong Kong's South China Morning Post was the first to use its own website and the first in Asia, having launched on April 19, 2005
- Communication from space. On 7 August 2005. American astronaut Steve Robinson claimed the first podcast from space during the Space Shuttle Discovery mission STS-114 - although there was no subscription feed, merely an audio file that required manual downloading. (transcript & audio).
- Conference and meeting alerts. Podcasts can be packaged to alert attendees to agendas, hosted roundtables and daily feedback.

Music

Replacement for live music audio streams. Whereas streaming a performance live over the Internet requires careful coordination of person and machine, podcasting offers the ability to do slight time-shifting of performances and greatly reduces the complexity of the effort. The quality of the program is often higher as post-production adjustments can be made prior to release. For example, programs such as the Woodsongs Old Time Radio Hour provide a live stream of their program, but most listeners don't hear it until weeks later on NPR. Podcasted versions of the programs split the difference, usually coming out a few days after the live program, but well before the traditional broadcast.

Politics

Politics: In the U.S., both major political parties have various podcasts, as do numerous politicians.

Publicity & Marketing

As a promotional vehicle for an upcoming event, such as Pixar's Cars Video Podcast, which advertised the release of Disney/Pixar's Cars animated feature film with a series of behind-the-scenes clips.

Health

Health, fitness and wellness resources, both general, such as Health Matters, and Family health Radio, and specific, such as AA Just For Today which covers self-rehabilitation from alcoholism, and Inner Path Yoga a fitness podcast.

Special Interests

Farm Podcasting makes information available about farming. The term was coined to identify a program that is produced exclusively as a podcast on the subject of agriculture. There are now multiple companies who specialize in farm podcasting and are producing regular programming targeted to farmers and the general public on the subject of agriculture.

Non-Traditional & Alternative Content

A way for people and organizations to avoid regulatory bodies, such as the British Ofcom, or American FCC that would not allow a program to be broadcast in traditional media.

Data & Networking Practises – Computer Network

What Is It?

A computer network is two or more computers connected together using a telecommunication system for the purpose of communicating and sharing resources.

Experts in the field of networking debate whether two computers that are connected together using some form of communications medium constitute a network. Therefore, some works state that a network requires three connected computers. For example, "Telecommunications: Glossary of Telecommunication Terms" states that a computer network is "A network of data processing nodes that are interconnected for the purpose of data communication", the term "network" being defined in the same document as "An interconnection of three or more communicating entities". A computer connected to a non-computing device (e.g., networked to a printer via an Ethernet link) may also represent a computer network, although this article does not address this configuration.

This article uses the definition, which requires two or more computers to be connected together to form a network. The same basic functions are generally present in this case as with larger numbers of connected computers.

Basic Components Of Computer Networks

Computers

Many of the components of an average network are individual computers, which are generally either workstations (including personal computers) or servers.

Computers: Types of workstations

There are many types of workstations that may be incorporated into a particular network, some of which have high-end displays, multiple CPUs, large amounts of RAM, large amounts of hard drive storage space, or other enhancements required for special data processing tasks, graphics, or other resource intensive applications. (See also network computer).

Computers: Types of servers

The following lists some common types of servers and their purpose.

- File server – Stores various types of files and distributes them to other clients on the network.
- Print server – Controls and manages one or more printers and accepts print jobs from other network clients, spooling the print jobs, and performing most or all of the other functions that a workstation would perform to accomplish a printing task if the printer were connected directly to the workstation's printer port.
- Mail server – Stores, sends, receives, routes, and performs other email related operations for other clients on the network.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

- Fax server – Stores, sends, receives, routes, and performs other functions necessary for the proper transmission, reception, and distribution of faxes.

- Telephony server – Performs telephony related functions such as answering calls automatically, performing the functions of an interactive voice response system, storing and serving voice mail, routing calls between the Public Switched Telephone Network (PSTN) and the network or the Internet (e.g., voice over IP (VoIP) gateway), etc.

- Proxy server – Performs some type of function on behalf of other clients on the network to increase the performance of certain operations (e.g., pre-fetching and caching documents or other data that are requested very frequently) or as a security precaution to isolate network clients from external threats.

- Remote Access Server (RAS) – Monitors modem lines or other network communications channels for requests to connect to the network from a remote location, answers the incoming telephone call or acknowledges the network request, and performs the necessary security checks and other procedures necessary to log a user onto the network.

- Application server – Performs the data processing or business logic portion of a client application, accepting instructions for operations to perform from a workstation and serving the results back to the workstation, while the workstation performs the user interface or GUI portion of the processing (i.e., the presentation logic) that is required for the application to work properly.

- Web server – Stores HTML documents, images, text files, scripts, and other Web related data (collectively known as content), and distributes this content to other clients on the network on request.

- Backup server – Has network backup software installed and has large amounts of hard drive storage or other forms of storage (tape, etc.) available to it to be used for the purpose of ensuring that data loss does not occur in the network.

Printers

Many printers are capable of acting as part of a computer network without any other device, such as a print server, to act as an intermediary between the printer and the device that is requesting a print job to be completed.

Thin Clients

Many networks use thin clients instead of workstations either for data entry and display purposes or in some cases where the application runs entirely on the server.

Other Devices

There are many other types of devices that may be used to build a network, many of which require an understanding of more advanced computer networking concepts before they are able to be easily understood (e.g., hubs, routers, bridges, switches, hardware firewalls, etc.). On home and mobile networks, connecting consumer electronics devices such as video game consoles is becoming increasingly common.

Building A Computer Network

A Simple Network

A simple computer network may be constructed from two computers by adding a network adapter (Network Interface Controller (NIC)) to each computer and then connecting them together with a special cable called a crossover cable. This type of network is useful for transferring information between two computers that are not normally connected to each other by a permanent network connection or for basic home networking applications. Alternatively, a network between two computers can be established without dedicated extra hardware by using a standard connection such as the RS-232 serial port on both computers, connecting them to each other via a special cross linked null modem cable.

Practical Networks

Practical networks generally consist of more than two interconnected computers and generally require special devices in addition to the Network Interface Controller that each computer needs to be equipped with.

Types Of Networks

Below is a list of the most common types of computer networks.

Local Area Network (LAN)

A network that is limited to a relatively small spatial area such as a room, a single building, a ship, or an aircraft. Local area networks are sometimes called a single location network.

Note: For administrative purposes, large LANs are generally divided into smaller logical segments called workgroups. A workgroup is a group of computers that share a common set of resources within a LAN.

Campus Area Network (CAN)

A network that connects two or more LANs but that is limited to a specific (possibly private) geographical area such as a college campus, industrial complex, or a military base

Note: A CAN is generally limited to an area that is smaller than a Metropolitan Area Network.

Metropolitan Area Network (MAN)

A network that connects two or more Local Area Networks or CANs together but does not extend beyond the boundaries of the immediate town, city, or metropolitan area. Multiple routers, switches & hubs are connected to create a MAN.

Wide Area Networks (WAN)

A WAN is a data communications network that covers a relatively broad geographic area and that often uses transmission facilities provided by common carriers, such as telephone companies. WAN technologies generally function at the lower three layers of the OSI reference model: the physical layer, the data link layer, and the network layer.

Types of WANs:

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

- Centralized – A centralized WAN consists of a central computer that is connected to dumb terminals and / or other types of terminal devices.
- Distributed – A distributed WAN consists of two or more computers in different locations and may also include connections to dumb terminals and other types of terminal devices.

Internetwork

Two or more networks or network segments connected using devices that operate at layer 3 (the 'network' layer) of the OSI Basic Reference Model, such as a router.

Note: Any interconnection among or between public, private, commercial, industrial, or governmental networks may also be defined as an internetwork.

Internet, The

A specific internetwork, consisting of a worldwide interconnection of governmental, academic, public, and private networks based upon the Advanced Research Projects Agency Network (ARPANET) developed by ARPA of the U.S. Department of Defense – also home to the World Wide Web (WWW) and referred to as the 'Internet' with a capital 'I' to distinguish it from other generic internetworks.

Intranet

A network or internetwork that is limited in scope to a single organization or entity or, also, a network or internetwork that is limited in scope to a single organization or entity and which uses the TCP/IP protocol suite, HTTP, FTP, and other network protocols and software commonly used on the Internet.

Note: Intranets may also be categorized as a LAN, CAN, MAN, WAN, or other type of network.

Extranet

A network or internetwork that is limited in scope to a single organization or entity but which also has limited connections to the networks of one or more other usually, but not necessarily, trusted organizations or entities (e.g., a company's customers may be provided access to some part of its intranet thusly creating an extranet while at the same time the customers may not be considered 'trusted' from a security standpoint).

Note: Technically, an extranet may also be categorized as a CAN, MAN, WAN, or other type of network, although, by definition, an extranet cannot consist of a single LAN, because an extranet must have at least one connection with an outside network.

Intranets and extranets may or may not have connections to the Internet. If connected to the Internet, the intranet or extranet is normally protected from being accessed from the Internet without proper authorization. The Internet itself is not considered to be a part of the intranet or extranet, although the Internet may serve as a portal for access to portions of an extranet.

Classification Of Computer Networks

By Network Layer

Computer networks may be classified according to the network layer at which they operate according to some basic reference models that are considered to be standards in the industry such as the seven layer OSI reference model and the five layer TCP/IP model.

By Scale

Computer networks may be classified according to the scale or extent of reach of the network, for example as a Personal area network (PAN), Local area network (LAN), Campus area network (CAN), Metropolitan area network (MAN), or Wide area network (WAN).

By Connection Method

Computer networks may be classified according to the technology that is used to connect the individual devices in the network such as HomePNA, Power line communication, Ethernet, or Wireless LAN.

By Functional Relationship

Computer networks may be classified according to the functional relationships which exist between the elements of the network, for example Active Networking, Client-server and Peer-to-peer (workgroup) architectures. Also, computer networks are used to send data from one to another by the hard drive.

By Network Topology

Computer networks may be classified according to the network topology upon which the network is based, such as Bus network, Star network, Ring network, Mesh network, Star-bus network, Tree or Hierarchical topology network, etc.

Topology can be arranged in a Geometric Arrangement

By Services Provided

Computer networks may be classified according to the services which they provide, such as Storage area networks, Server farms, Process control networks, Value-added network, Wireless community network, etc.

By Protocol

Computer networks may be classified according to the communications protocol that is being used on the network. See the articles on List of network protocol stacks and List of network protocols for more information.

Data & Networking Practises – Network Protocol

What Is It?

In computing, a protocol is a convention or standard that controls or enables the connection, communication, and data transfer between two computing endpoints. In its simplest form, a protocol can be defined as the rules governing the syntax, semantics, and synchronization of communication. Protocols may be implemented by hardware, software, or a combination of the two. At the lowest level, a protocol defines the behaviour of a hardware connection.

Typical Properties

It is difficult to generalize about protocols because they vary so greatly in purpose and sophistication. Most protocols specify one or more of the following properties:

- Detection of the underlying physical connection (wired or wireless), or the existence of the other endpoint or node
- Handshaking
- Negotiation of various connection characteristics
- How to start and end a message
- How to format a message
- What to do with corrupted or improperly formatted messages (error correction)
- How to detect unexpected loss of the connection, and what to do next
- Termination of the session or connection.

Importance

The widespread use and expansion of communications protocols is both a prerequisite to the Internet, and a major contributor to its power and success. The pair of Internet Protocol (or IP) and Transmission Control Protocol (or TCP) are the most important of these, and the term TCP/IP refers to a collection (or protocol suite) of its most used protocols. Most of the Internet's communication protocols are described in the RFC documents of the Internet Engineering Task Force (or IETF).

Object-oriented programming has extended the use of the term to include the programming protocols available for connections and communication between objects.

Generally, only the simplest protocols are used alone. Most protocols, especially in the context of communications or networking, are layered together into protocol stacks where the various tasks listed above are divided among different protocols in the stack.

Whereas the protocol stack denotes a specific combination of protocols that work together, the Reference Model is a software architecture that lists each layer and the services each should offer. The classic seven-layer reference model is the OSI model, which is used for conceptualising protocol stacks and peer entities. This reference model also provides an opportunity to teach more general software engineering concepts like hiding, modularity, and delegation of tasks. This model has endured in spite of the demise of many of its protocols (and protocol stacks) originally sanctioned by the ISO. The ISO model is not the only reference model however.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www*.planetlimited.*io

Common Protocols

- HTTP (Hyper Text Transfer Protocol)
- POP3 (Post Office Protocol 3).
- SMTP (Simple Mail Transfer Protocol).
- FTP (File Transfer Protocol).
- IP (Internet Protocol).
- DHCP (Dynamic Host Configuration Protocol).
- IMAP (Internet Message Access Protocol).

Data & Networking Practises – Ambient Network

What Is It?

Ambient Networks is a network integration solution to the modern-day problems of switching from one network to the other in order to keep in contact with the outside world. This project aims to develop a network software-driven infrastructure that will run on top of all current or future network physical infrastructures to provide a way for devices to connect to each other, and through each other to the outside world.

The concept of Ambient Networks comes from the IST Ambient Network project, which is a research project sponsored by the European Commission within the Sixth Framework Programme (FP6).

The Ambient Networks Project

Ambient Networks is a large-scale collaborative project within the European Union's Sixth Framework Program that investigates future communications systems beyond today's fixed and 3rd generation mobile networks. It is part of the Wireless World Initiative. The project works at a new concept called Ambient Networking, to provide suitable mobile networking technology for the future mobile and wireless communications environment.

Ambient Networks aims to provide a unified networking concept that can adapt to the very heterogeneous environment of different radio technologies and service and network environments.

Special focus is put on facilitating both competition and cooperation of various market players by defining interfaces, which allow the instant negotiation of agreements. This approach goes clearly beyond interworking of well-defined protocols and is expected to have a long-term effect on the business landscape in the Wireless World. Central to the project is the concept of composition of networks, which is an approach to address the dynamic nature of the target environment. The approach is based on an open framework for network control functionality, which can be extended with new capabilities as well as operating over existing connectivity infrastructure.

- Phase 1 of the project (2004-2005) has laid the conceptual foundations. The Deliverable D1-5 "Ambient Networks Framework Architecture" summarizes the work from phase 1 and provides links to other relevant material. (dead link)
- Ambient Networks Phase 2 (2006-2007) focuses on validation aspects. One key result of phase 2 is an integrated prototype that will be used to study the feasibility of the Ambient Networks concept for a number of typical network scenarios. The ACS prototype will be used to iteratively test the components developed by the project in a real implementation. In parallel, the top-down work is being continued which will lead to a refined System Specification. Furthermore, standardization of the composition concept is addressed in 3GPP.

Interfaces & Their Use

The ACS (Ambient Control Space) is the internal of an Ambient Network. It has the functions that can be accessed and it is in full control of the resources of the network. The Ambient Networks infrasturcture does not deal with nodes, instead it deals with networks, though at the beginning, all the "networks" might only consist of just one node: these "networks" need to merge in order to form a network in the original sense of the word. A composition establishment consists of the negotiation and then the realization of a Composition Agreement. This merging can happen be fully automatic. The decision to merge or not is decided using pre-configured policies.

There are three interfaces present to communicate with an ACS. These are:

- ANI: Ambient Network Interface. If a network wants to join in, it has to do so through this interface.
- ASI: Ambient Service Interface. If a function needs to be accessed inside the ACS, this Interface is used.
- ARI: Ambient Resource Interface. If a resource inside a network needs to be accessed (e.g. the volume of the traffic), this interface is used.

Interfaces are used in order to hide the internal structures of the underlying network.

If two networks meet, and decide to merge, a new ACS will be formed of the two (though the two networks will have their own ACS along with the interfaces inside this global, new ACS). The newly composed ACS will of course have its own ANI, ASI and ARI, and will use these interfaces in order to merge with other Ambient Networks. Other options for composition are to not merge the two Ambient Networks (Network Interworking) or to establish a new virtual ACS that exercises joint control over a given set of shared resources (Control Sharing).

ACS Functional Entities

Functions are divided into Functional Entities (FEs). The ACS provides a flexible and extensible framework to run these FEs as a distributed system. Examples are

- Composition Functional Entity: Controlling composition of ANs
- Bearer Management FEs
- Overlay Management FEs

Example Situation

Alice has a PAN, a Personal Area Network on her body: she has a Bluetooth enabled PDA, mobile phone and laptop that she is carrying, and are all currently turned on, and forming a network. Her laptop also has the ability to connect using an available WLAN, and her mobile phone has the ability to connect through GPRS, though GPRS is slower and much more costly for Alice to use. She is now on the move, and her laptop is downloading her emails using the GPRS connection on the mobile:

Laptop -> (Bluetooth) -> Mobile -> (GPRS) -> Mobile phone network

While walking, she passes into an area covered by a free WLAN hotspot: Her PAN now immediately starts to initiate a connection with the hotspot. This is called "merging" of the networks (that of the hotspot and that of her PAN). Once this merging is complete, the downloading of her email continues totally unaffected, but instead of using the expensive and slow GPRS connection, it is now using the newly established WLAN connection. If she now wants to browse the web with her PDA, the PDA will also use the WLAN connection of the laptop:

PDA-> (bluetooth) -> Laptop-> (WLAN) -> Hotspot

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

Data & Networking Practises – Replication

What Is It?

Replication refers to the use of redundant resources, such as software or hardware components, to improve reliability, fault-tolerance, or performance. Replication typically involves replication in space, in which the same data is stored on multiple storage devices or the same computing task is executed on multiple devices, or replication in time, in which a computing task is executed repeatedly on a single device.

Replication In Distributed Systems

There are two approaches to replication in distributed systems, active and passive replication. active replication, also known as state machine replication, is performed by processing the same request at every replica. In passive replication, requests are usually processed on a single replica and then the state is transferred to the other replicas. If there is only one machine that processes the requests, then we are talking about the primary-backup scheme. On the other side, if any machine can process a request, then we have a multi-primary scheme. In the multi-primary scheme, some form of distributed concurrency control must be used.

Database Replication

Database replication can be used on many database management systems, usually with a master/slave relationship between the original and the copies. The master logs the updates, which then ripple through to the slaves. The slave outputs a message stating that it has received the update successfully, thus allowing the sending (and potentially re-sending until successfully applied) of subsequent updates. See also Coda and RAID. Multi-master replication, where updates can be submitted to any database node, and then "ripple" through to other servers, is often desired, but introduces substantially increased costs and complexity which may make it impractical in some situations.

The most common challenge that exists in multi-master replication is conflict resolution. For instance, if records are changed in two systems simultaneously, the resolution of that conflict can take many paths. One simple method is that of timing, where data with the first timestamp wins. Alternately, data with the latest timestamp could be saved as most valuable. Another way of resolving conflicts is through hierarchical rules, having declared sites and/or users to have greater rights that supersede changes of lower sites/users. Finally, logic-based conflict resolution can be employed, which is more configurable, but more complex.

Filesystem Replication

Active (real-time) file system replication is usually implemented by distributing updates of a virtual block device to several physical hard disks. This way, any filesystem supported by the operating system can be replicated without modification, as the file system code works on a level above the block device layer. The most popular method for filesystem replication is RAID which is typically limited to locally-connected disks only.

Alternatively, updates to a block device can be replicated (that is, distributed) over a computer network. This has the advantage that the replication slaves can be located in physically distant locations, to avoid damage done by, and improve availability in case of local failures or disasters. An example of this kind of replication is the DRBD module for Linux.

Distributed Shared Memory Replication

Another example of using replication appears in distributed shared memory systems, where it may happen that many nodes of the system share the same page of the memory - which usually means, that each node has a separate copy (replica) of this page.

Replication Transparency

If a resource is replicated among several locations, it should appear to the user as a single resource.

Data Replication

Data replication refers to a data storage and backup strategy that copies data from a host computer to another computer, which may or may not be at a remote location. Replication over a computer network can make data backup entirely independent of local data center physical storage.

Replication is popular in Relational Database Management Systems where it can be used to provide redundancy or to balance load across multiple database servers.

Online Data Replication

In data replication over a network such as the Internet, the data backup is made on a real time basis wherein the data from the host server is copied onto the remote location as soon as the data is changed.

Offline Data Replication

In offline data replication, a backup of the datafile is taken on a remote location on an offline basis, such as once a day. Offline data replication is preferred in environments where the number of transactions is very large and doing real time data replication may affect the performance of the system.

Active/Active Replication

Traditional approaches to replication are based on a master/slave model where one device or process has unilateral control over one or more other devices. This approach has many flaws, not least, that only one process may be active at any one time, which means that only one user could edit a unit of information at any one time. The other approach is multi-master technology. The problem here is that it can not efficiently operate in a WAN and can only efficiently operate in a Local Area Network (LAN).

WANdisco has developed a new mathematical theory that enables active/active replication where every node on a network is an exact copy or replica and hence every node on the network is active at one time. Hence there is no master/slave paradigm. WANdisco is able to achieve this replication over a wide area network (WAN).

Data & Networking Practises – File Sharing

What Is It?

File sharing is the practice of making files available for other users to download over the Internet and smaller networks. Usually file sharing follows the peer-to-peer (P2P) model, where the files are stored on and served by personal computers of the users. Most people who engage in file sharing are also downloading files that other users share. Sometimes these two activities are linked together. P2P file sharing is distinct from file trading in that downloading files from a P2P network does not require uploading, although some networks either provide incentives for uploading such as credits or force the sharing of files being currently downloaded.

The First P2P-Generation: Server-Client

The first generation of peer-to-peer file sharing networks had a centralized file list. Courts in the United States ruled that whoever controlled this centralized file list, containing works whose copyright was being infringed upon, were responsible for any infringement. Ultimately, Napster was held liable even if it used the most advanced technology available to identify works copyright holders had asked it to block, because no technology that can identify works with 100% certainty exists or can exist. Napster still exists today, but as a subsidiary of Roxio after they bought the name during the original Napster's bankruptcy phase, continues to operate today, and is now legally distributing music under a subscription-based model.

In the centralized peer-to-peer model, a user would send a search to the centralized server of what they were looking for, that is, song, video, movie. The server then sends back a list of which peers have the data and facilitates the connection and download.

The first file sharing programs marked themselves by inquiries to a server, either the data to the Download held ready or in appropriate different Peers and so-called Nodes further-obtained, so that one could download there. Best example was Napster (today a Payment offerer) or eDonkey2000 in the server version (today likewise with Overnet and KAD -network decentralized).

The Second P2P-Generation: Decentralization

After Napster encountered legal troubles, Justin Frankel of Nullsoft set out to create a network without a central index server, and Gnutella was the result. Unfortunately, the Gnutella model of all nodes being equal quickly died from bottlenecks as the network grew from incoming Napster refugees. FastTrack solved this problem by having some nodes be 'more equal than others'.

By electing some higher capacity nodes to be indexing nodes, with lower capacity nodes branching off from them, FastTrack allowed for a network that could scale to a much larger size. Gnutella quickly adopted this model, and most current peer-to-peer networks implement this design, as it allows for large and efficient networks without central servers.

Also included in the second generation are distributed hash tables (DHTs), which help solve the scalability problem by electing various nodes to index certain hashes (which are used to identify files), allowing for fast and efficient searching for any instances of a file on the

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited.**io

network. This is not without drawbacks; perhaps most significantly, DHTs do not directly support keyword searching (as opposed to exact-match searching).

The best examples are Gnutella, Kazaa or eMule with Kademlia, whereby Kazaa has still a central server for logging in. eDonkey2000/Overnet, Gnutella, FastTrack and Ares Galaxy have summed up approx. 10.3 million users (as of April 2006, according to slyck.com). This number not necessarily corresponds to the actual number of persons who use these networks; it must be assumed that some use multiple clients for different networks. The number of BitTorrent users cannot be measured directly, however, there should be more users than of eDonkey2000 and that's why providers try to restrict the BitTorrent traffic first. One must also understand that number of users indicated by software applies only to users which are active at this moment.

The Third P2P-Generation: Non-Direct & Encrypted

The third generation of peer-to-peer networks are those that have anonymity features built in. Examples of anonymous networks are ANts P2P, RShare, Freenet, I2P, GNUnet and Entropy.

A degree of anonymity is realized by routing traffic through other users clients, which have the function of network nodes.This makes it harder for someone to identify who is downloading or who is offering files. Most of these programs also have strong encryption to resist traffic sniffing.

Friend-to-friend networks only allow already known users (also known as "friends") to connect to your computer, then each node can forward requests and files anonymously between its own "friends" nodes.

Third generation networks have not reached mass usage for file sharing because most current implementations incur too much overhead in their anonymity features, making them slow or hard to use.

An example might be: Petra gives a file to Oliver, then Oliver gives the file to Anna. Petra and Anna thus never become acquainted and thus are protected. Virtual IP addresses are also often used, further obfuscating the user's network location. Additionally all transfers are encrypted, so that even the network administrators cannot see what was sent to whom. Example software includes WASTE, JetiANts, Tor and I2P. These clients differ greatly in their goals and implementation. WASTE is designed only for small groups and may therefore be considered Darknet, ANts and I2P are public Peer to Peer systems, with anonymization provided exclusively by routing reach.

The Fourth P2P-Generation: Streams Over P2P

Apart from the traditional file sharing there are services that send streams instead of files over a P2P network. Thus one can hear radio and watch television without any server involved -- the streaming media is distributed over a P2P network. It is important that instead of a treelike network structure, a swarming technology known from BitTorrent is used. Best examples are Peercast, Cybersky and demo TV.

- General examples include Broadcatching and Podcasting
- Tree structure examples include Peercast and CoolStreaming
- Swarm structure such as Bittorrent, TVUPlayer, Joost, PPLive, Icecast, PeerCast, PPStream, SopCast, MediaBlog and Zudeo

Copyright issues

File sharing (such as with the Gnutella and Napster networks) grew in popularity with the proliferation of high speed Internet connections, relatively small file size and high-quality MP3 audio format. Although file sharing is a legal technology with legal uses, many users use it to download and upload copyrighted materials without permission, which can be copyright infringement if done without authorization for improper purposes. This has led to attacks against file sharing in general from many copyright owners.

There has been great discussion over perceived and actual legal issues surrounding file sharing. In circumstances where trading partners are in different countries with different legal codes, there are significant problems to contend with. What if a person in Canada wishes to share a piece of source code which, if compiled, has encryption capabilities? In some countries, a citizen may not request or receive such information without special permission.

Throughout the early 2000s, the entire file-sharing community has been in a state of flux. In the year 2000, there was speculation over how seriously record companies and the Recording Industry Association of America would strike the file-sharing community because of its limits compared to more traditional forms of media. However, the communities suffered strain as record companies and the RIAA tried to shut down as much of it as possible. Even though they have forced Napster and Grokster into cooperating against copyright violations, they are fighting an uphill battle since the community has flourished and produced many different clients based on several different underlying protocols. The third generation of P2P protocols, such as Freenet, are not as dependent as Napster is on a central server; and as they encrypt the shared data, it is much harder to shut down these systems through court actions. Another attempt (used by the maintainers of KaZaA) is to change the company's organization or country of origin so that it is impossible or useless to attack it legally. To date, file sharing in Canada is sometimes legal (unauthorized downloading of music is permitted, for example) , though not completely so. The uphill battle also extends to the legal actions taken by the RIAA and motion picture counterpart MPAA against individuals using file-sharing programs to distribute material protected by copyright. Ambiguity in the interpretation of copyright law has been a major factor contributing the lack of successful enforcement by the Intellectual Property owners. In Electra v. Perez for example, the Court ruled that the act of making such files available for distribution equated to infringement of the works involved. In the more recent UMG v. Lindor, Judge David G. Trager ruled that the RIAA would be required to prove that actual distribution (sharing) occurred. Should this case go to trial in 2007, it will likely set a precedent for further RIAA and MPAA actions.

The Electronic Frontier Foundation (EFF) is a donor-supported group which seeks to protect and expand digital rights. Its activities include litigation, political lobbying, and public awareness campaigns. The EFF has vocally opposed the RIAA in its pursuit of lawsuits against users of file sharing applications and supported defendants in these cases. The foundation promotes the legalization of peer-to-peer sharing of copyrighted materials and alternative methods to provide compensation to copyright holders.

Risks

Some file sharing software comes bundled with malware such as spyware or adware. Sometimes this malware remains installed on the system even if the original file sharing software is removed, and can be very difficult to eliminate. In many cases such malware can interfere with the correct operation of web browsers, anti-virus software, anti-spyware and software firewalls, and can cause degraded performance on affected systems. Such malware is typically bundled with proprietary software, and not those in open source. In most cases it is possible to easily remove adware and spyware by running spyware removal programs. Such programs can often remove malware without influencing the functionality of the file sharing software.

Some are also concerned about the use of file sharing systems to distribute child pornography, inflammatory literature, and illegal or unpopular material. Novice users may find it difficult to obtain information on which networks are "safe" for them to use. However, experienced users know that there is only one way to get in contact with such material: You have to actively search for it. Therefore they recommend not to search for illegal material.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

Data & Networking Practises – Distributed Computing

What Is It?

Distributed computing is a method of computer processing in which different parts of a program run simultaneously on two or more computers that are communicating with each other over a network. Distributed computing is a type of parallel processing. But the latter term is most commonly used to refer to processing in which different parts of a program run simultaneously on two or more processors that are part of the same computer. While both types of processing require that a program be parallelized -- divided into sections that can run simultaneously, distributed computing also requires that the division of the program take into account the different environments on which the different sections of the program will be running. For example, two computers are likely to have different file systems and different hardware components.

An example of distributed computing is BOINC, a service in which large problems can be divided into many small problems which are distributed to many computers. Later, the small results are reassembled into a larger solution.

Distributed computing is a natural result of the use of networks to allow computers to efficiently communicate. But distributed computing is distinct from networking. The latter refers to two or more computers interacting with each other, but not, typically, sharing the processing of a single program. The World Wide Web is an example of a network, but not an example of distributed computing.

There are numerous technologies and standards used to construct distributed computations, including some which are specially designed and optimized for that purpose, such as Remote Procedure Calls (RPC) or Remote Method Invocation (RMI) or .NET Remoting.

Organization

Organizing the interaction between each computer is of prime importance. In order to be able to use the widest possible range and types of computers, the protocol or communication channel should not contain or use any information that may not be understood by certain machines. Special care must also be taken that messages are indeed delivered correctly and that invalid messages are rejected which would otherwise bring down the system and perhaps the rest of the network.

Another important factor is the ability to send software to another computer in a portable way so that it may execute and interact with the existing network. This may not always be possible or practical when using differing hardware and resources, in which case other methods must be used such as cross-compiling or manually porting this software.

Goals & Advantages

There are many different types of distributed computing systems and many challenges to overcome in successfully designing one. The main goal of a distributed computing system is to connect users and resources in a transparent, open, and scalable way. Ideally this arrangement is drastically more fault tolerant and more powerful than many combinations of stand-alone computer systems.

Openness

Openness is the property of distributed systems such that each subsystem is continually open to interaction with other systems (see references). Web Services protocols are standards which enable distributed systems to be extended and scaled. In general, an open system that scales has an advantage over a perfectly closed and self-contained system.

Consequently, open distributed systems are required to meet the following challenges:

- Monotonicity – Once something is published in an open system, it cannot be taken back.
- Pluralism – Different subsystems of an open distributed system include heterogeneous, overlapping and possibly conflicting information. There is no central arbiter of truth in open distributed systems.
- Unbounded nondeterminism – Asynchronously, different subsystems can come up and go down and communication links can come in and go out between subsystems of an open distributed system. Therefore the time that it will take to complete an operation cannot be bounded in advance (see unbounded nondeterminism).

Scalability

A scalable system is one that can easily be altered to accommodate changes in the number of users, resources and computing entities affected to it. Scalability can be measured in three different dimensions:

- Load scalability – A distributed system should make it easy for us to expand and contract its resource pool to accommodate heavier or lighter loads.
- Geographic scalability – A geographically scalable system is one that maintains its usefulness and usability, regardless of how far apart its users or resources are.
- Administrative scalability – No matter how many different organizations need to share a single distributed system, it should still be easy to use and manage.

Some loss of performance may occur in a system that allows itself to scale in one or more of these dimensions. There is a limit up to which we can scale/add processors to the system, and above that the performance of the system degrades.

Drawbacks & Disadvantages

If not planned properly, a distributed system can decrease the overall reliability of computations if the unavailability of a node can cause a disruption of the other nodes. Leslie Lamport describes this type of distributed system fragility like this: "You know you have one when the crash of a computer you've never heard of stops you from getting any work done."

Troubleshooting and diagnosing problems in a distributed system can also become more difficult, because the analysis may now require connecting to remote nodes or inspecting communications being sent between nodes.

Not many types of computation are well-suited for distributed environments, due typically to the amount of network communication or synchronization that would be required between nodes. If bandwidth, latency, or communication requirements are too significant, then the benefits of distributed computing may be negated and the performance may be worse than a non-distributed environment.

Fallacies Of Distributed Computing

The Fallacies of Distributed Computing are a set of common but flawed assumptions made by programmers when first developing distributed applications. The fallacies are summarized as follows:

1. The network is reliable.
2. Latency is zero.
3. Bandwidth is infinite.
4. The network is secure.
5. Topology doesn't change.
6. There is one administrator.
7. Transport cost is zero.
8. The network is homogeneous.

The list of fallacies generally came about at Sun Microsystems. Peter Deutsch, one of the original Sun "Fellows," is credited with penning the first seven fallacies in 1994; however, Bill Joy and Tom Lyon had already identified the first four as "The Fallacies of Networked Computing" (the article claims "Dave Lyon," but this is considered a mistake). Around 1997, James Gosling, another Sun Fellow and the inventor of Java, added the eighth fallacy.

Architecture

Various hardware and software architectures are used for distributed computing. At a lower level, it is necessary to interconnect multiple CPUs with some sort of network, regardless of whether that network is printed onto a circuit board or made up of loosely coupled devices and cables. At a higher level, it is necessary to interconnect processes running on those CPUs with some sort of communication system.

Distributed programming typically falls into one of several basic architectures or categories: Client-server, 3-tier architecture, N-tier architecture, Distributed objects, loose coupling, or tight coupling.

- Client-server — Smart client code contacts the server for data, then formats and displays it to the user. Input at the client is committed back to the server when it represents a permanent change.
- 3-tier architecture — Three tier systems move the client intelligence to a middle tier so that stateless clients can be used. This simplifies application deployment. Most web applications are 3-Tier.
- N-tier architecture — N-Tier refers typically to web applications which further forward their requests to other enterprise services. This type of application is the one most responsible for the success of application servers.
- Tightly coupled (clustered) — refers typically to a set of highly integrated machines that run the same process in parallel, subdividing the task in parts that are made individually by each one, and then put back together to make the final result.
- Peer-to-peer — an architecture where there is no special machine or machines that provide a service or manage the network resources. Instead all responsibilities are uniformly divided among all machines, known as peers.

Concurrency

Distributed computing implements a kind of concurrency. It interrelates tightly with concurrent programming so much that they are sometimes not taught as distinct subjects.

Multiprocessor Systems

A multiprocessor system is simply a computer that has more than one CPU on its motherboard. If the operating system is built to take advantage of this, it can run different processes (or different threads belonging to the same process) on different CPUs.

Multicore Systems

Intel CPUs from the late Pentium 4 era (Northwood and Prescott cores) employed a technology called Hyperthreading that allowed more than one thread (usually two) to run on the same CPU. The most recent Sun UltraSPARC T1, AMD Athlon 64 X2, AMD Athlon FX, AMD Opteron, Intel Pentium D and Intel Core 2 processors feature multiple processor cores to also increase the number of concurrent threads they can run.

Multicomputer Systems

A multicomputer system is a system made up of several independent computers interconnected by a telecommunications network.

Multicomputer systems can be homogeneous or heterogeneous: A homogeneous distributed system is one where all CPUs are similar and are connected by a single type of network. They are often used for parallel computing.

A heterogeneous distributed system is made up of different kinds of computers, possibly with vastly differing memory sizes, processing power and even basic underlying architecture. They are in widespread use today, with many companies adopting this architecture due to the speed with which hardware goes obsolete and the cost of upgrading a whole system simultaneously.

Computing Taxonomies

The types of distributed computers are based on Flynn's taxonomy of systems; single instruction, single data (SISD), single instruction, multiple data (SIMD), multiple instruction, single data (MISD), and multiple instruction, multiple data (MIMD).

Computer Clusters

A cluster consists of multiple stand-alone machines acting in parallel across a local high speed network. Distributed computing differs from cluster computing in that computers in a distributed computing environment are typically not exclusively running "group" tasks, whereas clustered computers are usually much more tightly coupled. Distributed computing also often consists of machines, which are widely separated geographically.

Grid Computing

A grid uses the resources of many separate computers connected by a network (usually the Internet) to solve large-scale computation problems. Most use idle time on many thousands of computers throughout the world. Such arrangements permit handling of data that would otherwise require the power of expensive supercomputers or would have been impossible to analyze.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

Languages

Nearly any programming language that has access to the full hardware of the system could handle distributed programming given enough time and code. Remote procedure calls distribute operating system commands over a network connection. Systems like CORBA, Microsoft D/COM, Java RMI and others, try to map object-oriented design to the network. Loosely coupled systems that communicate through intermediate documents that are typically human readable are XML, HTML, SGML, X.500, and EDI.

Languages specifically tailored for distributed programming are:

- Ada programming language
- Alef programming language
- E programming language
- Erlang programming language
- Limbo programming language
- Oz programming language

Examples

Projects

A variety of distributed computing projects have grown up in recent years. Many are run on a volunteer basis, and involve users donating their unused computational power to work on interesting computational problems. Examples of such projects include the Stanford University Chemistry Department Folding@home project, which is focused on simulations of protein folding to find disease cures; World Community Grid, an effort to create the world's largest public computing grid to tackle scientific research projects that benefit humanity, run and funded by IBM; SETI@home, which is focused on analyzing radio-telescope data to find evidence of intelligent signals from space, hosted by the Space Sciences Laboratory at the University of California, Berkeley; and distributed.net, which is focused on breaking various cryptographic ciphers.

Distributed computing projects also often involve competition with other distributed systems. This competition may be for prestige, or it may be a matter of enticing users to donate processing power to a specific project. For example, stat races are a measure of the work a distributed computing project has been able to compute over the past day or week. This has been found to be so important in practice that virtually all distributed computing projects offer online statistical analyses of their performances, updated at least daily if not in real-time.

World Wide Web

An example of a distributed system (not distributed computing) is the World Wide Web. As you are reading a web page, you are actually using the distributed system that comprises the site. As you are browsing the web, your web browser running on your own computer communicates with different web servers that provide web pages. Possibly, your browser uses a proxy server to access the web contents stored on web servers faster and more securely. To find these servers, it also uses the distributed domain name system. Your web browser communicates with all of these servers over the Internet, via a system of routers which are themselves part of a large distributed system.

Data & Networking Practises – Grid Computing

What Is It?

Grid computing is an emerging computing model that distributes processing across a parallel infrastructure. Throughput is increased by networking many heterogeneous resources across administrative boundaries to model a virtual computer architecture. For a computing problem to benefit from a grid, it must require either large amounts of computation time or large amounts of data, and it must be reducible to parallel processes that do not require intensive inter-communication. Today resource allocation in a grid is done in accordance with SLAs (service level agreements).

Origins

Like the Internet, grid computing evolved from the computational needs of "big science". The Internet was developed to meet the need for a common communication medium between large, federally funded computing centres. These communication links led to resource and information sharing between these centres and eventually to provide access to them for additional users. Ad hoc resource sharing 'procedures' among these original groups pointed the way toward standardization of the protocols needed to communicate between any administrative domain. The current grid technology can be viewed as an extension or application of this framework to create a more generic resource-sharing context.

Fully functional proto-grid systems date back to the early 1970s with the Distributed Computing System (DCS) project at the University of California, Irvine. David Farber was the main architect. This system was well known enough to merit coverage and a cartoon depiction in Business Week on 14 July 1973. The caption read "The ring acts as a single, highly flexible machine in which individual units can bid for jobs". In modern terminology ring = network, and units = computers, very similar to how computational capabilities are utilized on the grid. The project's final report was published in 1977. This technology was mostly abandoned in the 1980s as the administrative and security issues involved in having machines you did not control do your computation were (and are still by some) seen as insurmountable.

The ideas of the grid were brought together by Ian Foster, Carl Kesselman and Steve Tuecke, the so called "fathers of the grid." They led the effort to create the Globus Toolkit incorporating not just CPU management (examples: cluster management and cycle scavenging) but also storage management, security provisioning, data movement, monitoring and a toolkit for developing additional services based on the same infrastructure including agreement negotiation, notification mechanisms, trigger services and information aggregation. In short, the term grid has much further reaching implications than the general public believes. While Globus Toolkit remains the de facto standard for building grid solutions, a number of other tools have been built that answer some subset of services needed to create an enterprise grid.

Features

Grid computing offers a model for solving massive computational problems by making use of the unused resources (CPU cycles and/or disk storage) of large numbers of disparate computers, often desktop computers, treated as a virtual cluster embedded in a distributed telecommunications infrastructure. Grid computing's focus on the ability to support computation across administrative domains sets it apart from traditional computer clusters or traditional distributed computing.

Grids offer a way to solve Grand Challenge problems like protein folding, financial modelling, earthquake simulation, and climate/weather modeling. Grids offer a way of using the information technology resources optimally inside an organization. They also provide a means for offering information technology as a utility bureau for commercial and non-commercial clients, with those clients paying only for what they use, as with electricity or water.

Grid computing has the design goal of solving problems too big for any single supercomputer, whilst retaining the flexibility to work on multiple smaller problems. Thus Grid computing provides a multi-user environment. Its secondary aims are better exploitation of available computing power and catering for the intermittent demands of large computational exercises.

This approach implies the use of secure authorization techniques to allow remote users to control computing resources.

Grid computing involves sharing heterogeneous resources (based on different platforms, hardware/software architectures, and computer languages), located in different places belonging to different administrative domains over a network using open standards. In short, it involves virtualizing computing resources.

Grid computing is often confused with cluster computing. The key difference is that a cluster is a single set of nodes sitting in one location, while a Grid is composed of many clusters and other kinds of resources (e.g. networks, storage facilities).

Functionally, one can classify Grids into several types:

- Computational Grids (including CPU scavenging Grids) which focuses primarily on computationally-intensive operations.
- Data Grids or the controlled sharing and management of large amounts of distributed data.
- Equipment Grids which have a primary piece of equipment e.g. a telescope, and where the surrounding Grid is used to control the equipment remotely and to analyse the data produced.

Grid computing is presently being applied successfully by the National Science Foundation's National Technology Grid, NASA's Information Power Grid, Pratt & Whitney, Bristol-Myers Squibb, Co., and American Express.

Definitions

The term Grid computing originated in the early 1990s as a metaphor for making computer power as easy to access as an electric power grid.

Today there are many definitions of Grid computing:

- The definitive definition of a Grid is provided by Ian Foster in his article "What is the Grid? A Three Point Checklist". The three points of this checklist are:
    - Computing resources are not administered centrally.
    - Open standards are used.
    - Non-trivial quality of service is achieved.

- Plaszczak/Wellner define Grid technology as "the technology that enables resource virtualization, on-demand provisioning, and service (resource) sharing between organizations."
- IBM defines Grid Computing as "the ability, using a set of open standards and protocols, to gain access to applications and data, processing power, storage capacity and a vast array of other computing resources over the Internet. A Grid is a type of parallel and distributed system that enables the sharing, selection, and aggregation of resources distributed across 'multiple' administrative domains based on their (resources) availability, capacity, performance, cost and users' quality-of-service requirements"
- An earlier example of the notion of computing as utility was in 1965 by MIT's Fernando Corbató. Fernando and the other designers of the Multics operating system envisioned a computer facility operating "like a power company or water company". http://www.multicians.org/fjcc3.html
- Buyya defines Grid as "a type of parallel and distributed system that enables the sharing, selection, and aggregation of geographically distributed autonomous resources dynamically at runtime depending on their availability, capability, performance, cost, and users' quality-of-service requirements".
- CERN, one of the largest users of Grid technology, talk of The Grid: "a service for sharing computer power and data storage capacity over the Internet."
- Pragmatically, Grid computing is attractive to geographically-distributed non-profit collaborative research efforts like the NCSA Bioinformatics Grids such as BIRN: external Grids.
- Grid computing is also attractive to large commercial enterprises with complex computation problems who aim to fully exploit their internal computing power: internal Grids.
- A recent survey (done by Heinz Stockinger in spring 2006; to be published in the Journal of Supercomputing in early 2007) presents a snapshot on the view in 2006.
- Another survey (done by Miguel L. Bote-Lorenzo et al. in autumn 2002; published in the LNCS series of Springer-Verlag) presents a snapshot on the view in 2002.

Grids can be categorized with a three stage model of departmental Grids, enterprise Grids and global Grids. These correspond to a firm initially utilising resources within a single group i.e. an engineering department connecting desktop machines, clusters and equipment. This progresses to enterprise Grids where non-technical staff's computing resources can be used for cycle-stealing and storage. A global Grid is a connection of enterprise and departmental Grids which can be used in a commercial or collaborative manner.

Grid computing is a subset of distributed computing.

Conceptual Framework

Grid computing reflects a conceptual framework rather than a physical resource. The Grid approach is utilized to provision a computational task with administratively distant resources. The focus of Grid technology is associated with the issues and requirements of flexible computational provisioning beyond the local (home) administrative domain.

Virtual Organization

A Grid environment is created to address resource needs. The use of that resource(s) (eg. CPU cycles, disk storage, data, software programs, peripherals) is usually characterized by its availability outside of the context of the local administrative domain. This 'external provisioning' approach entails creating a new administrative domain referred to as a Virtual

Organization (VO) with a distinct and separate set of administrative policies (home administration policies plus external resource administrative policies equals the VO (aka your Grid) administrative policies). The context for a Grid 'job execution' is distinguished by the requirements created when operating outside of the home administrative context. Grid technology (aka. middleware) is employed to facilitate formalizing and complying with the Grid context associated with your application execution.

Resources

One characteristic that currently distinguishes Grid computing from distributed computing is the abstraction of a 'distributed resource' into a Grid resource. One result of abstraction is that it allows resource substitution to be more easily accomplished. Some of the overhead associated with this flexibility is reflected in the middleware layer and the temporal latency associated with the access of a Grid (or any distributed) resource. This overhead, especially the temporal latency, must be evaluated in terms of the impact on computational performance when a Grid resource is employed.

Web based resources or Web based resource access is an appealing approach to Grid resource provisioning. A recent GGF (Global Grid Forum) Grid middleware evolutionary development "re-factored" the architecture/design of the Grid resource concept to reflect using the W3C WSDL (Web Service Description Language) to implement the concept of a WS-Resource. The stateless nature of the Web, while enhancing the ability to scale, can be a concern for applications that migrate from a stateful protocol for accessing resources to the Web-based stateless protocol. The GGF WS-Resource concept includes discussions on accommodating the statelessness associated with Web resources access.

State-Of-The-Art

The conceptual framework and ancillary infrastructure are evolving at a fast pace and include international participation. The business sector is actively involved in commercialization of the Grid framework. The "big science" sector is actively addressing the development environment and resource (aka performance) monitoring aspects. Activity is also observed in providing Grid-enabled versions of HPC (High Performance Computing) tools. Activity in the domains of "little science" appears to be scant at this time. The treatment in the GGF documentation series reflects the HPC roots of the Grid concept framework; this bias should not be interpreted as a restriction in the application of the Grid conceptual framework in its application to other research domains or other computational contexts.

Substantial experience is being built through the operation of various Grids, most notable of them being the EGEE infrastructure supporting LCG, the Large Hadron Collider Computing Grid. LCG is driven by CERN's need to handle a huge amount of data, produced at a rate of almost a gigabyte per second (10 petabytes per year), a history not unlike that of the production NorduGrid. A list of active sites participating within LCG can be found online as can real time monitoring of the EGEE infrastructure. The relevant software and documentation is also publicly accessible.

Data & Networking Practises – P2P

What Is It?

A peer-to-peer (or P2P) computer network relies primarily on the computing power and bandwidth of the participants in the network rather than concentrating it in a relatively low number of servers. P2P networks are typically used for connecting nodes via largely ad hoc connections. Such networks are useful for many purposes. Sharing content files (see file sharing) containing audio, video, data or anything in digital format is very common, and realtime data, such as telephony traffic, is also passed using P2P technology.

A pure peer-to-peer network does not have the notion of clients or servers, but only equal peer nodes that simultaneously function as both "clients" and "servers" to the other nodes on the network. This model of network arrangement differs from the client-server model where communication is usually to and from a central server. A typical example for a non peer-to-peer file transfer is an FTP server where the client and server programs are quite distinct, and the clients initiate the download/uploads and the servers react to and satisfy these requests.

The earliest peer-to-peer network in widespread use was the Usenet news server system, in which peers communicated with one another in order to propagate Usenet news articles over the entire Usenet network. Particularly in the earlier days of Usenet, UUCP was used to extend even beyond the Internet. However, the news server system also acted in a client-server form when individual users accessed a local news server in order to read and post articles. The same consideration applies to SMTP email in the sense that the core email relaying network of Mail transfer agents is a peer-to-peer network while the periphery of Mail user agents and their direct connections is client server.

Some networks and channels such as Napster, OpenNAP and IRC server channels use a client-server structure for some tasks (e.g. searching) and a peer-to-peer structure for others. Networks such as Gnutella or Freenet use a peer-to-peer structure for all purposes, and are sometimes referred to as true peer-to-peer networks, although Gnutella is greatly facilitated by directory servers that inform peers of the network addresses of other peers.

Peer-to-peer architecture embodies one of the key technical concepts of the internet, described in the first internet Request for Comments, RFC 1, "Host Software" dated 7 April 1969. More recently, the concept has achieved recognition in the general public in the context of the absence of central indexing servers in architectures used for exchanging multimedia files.

The concept of peer to peer is increasingly evolving to an expanded usage as the relational dynamic active in distributed networks, i.e. not just computer to computer, but human to human. Yochai Benkler has coined the term "commons-based peer production" to denote collaborative projects such as free software. Associated with peer production are the concept of peer governance (referring to the manner in which peer production projects are managed) and peer property (referring to the new type of licenses which recognize individual authorship but not exclusive property rights, such as the GNU General Public License and the Creative Commons License).

Classification Of Peer-To-Peer Networks

One possible classification of peer-to-peer networks is according to their degree of centralization:

Pure peer-to-peer:

- Peers act as equals, merging the roles of clients and server
- There is no central server managing the network
- There is no central router

Hybrid peer-to-peer:

- Has a central server that keeps information on peers and responds to requests for that information.
- Peers are responsible for hosting available resources (as the central server does not have them), for letting the central server know what resources they want to share, and for making its shareable resources available to peers that request it.
- Route terminals are used addresses, which are referenced by a set of indices to obtain an absolute address.

Some examples of "pure" peer-to-peer application layer networks designed for file sharing are Gnutella and Freenet.

Meanwhile some may also categorize peer-to-peer networks into the following categories:

- Centralized P2P network such as Napster
- Decentralized P2P network such as KaZaA
- Structured P2P network such as CAN
- Unstructured P2P network such as Gnutella
- Hybrid P2P network (Centralized and Decentralized) such as JXTA

Some also categorize peer-to-peer networks into the following categories:

- 1st Generation P2P
- 2nd Generation P2P

Advantages Of Peer-To-Peer Networks

An important goal in peer-to-peer networks is that all clients provide resources, including bandwidth, storage space, and computing power. Thus, as nodes arrive and demand on the system increases, the total capacity of the system also increases. This is not true of a client-server architecture with a fixed set of servers, in which adding more clients could mean slower data transfer for all users.

The distributed nature of peer-to-peer networks also increases robustness in case of failures by replicating data over multiple peers, and -- in pure P2P systems -- by enabling peers to find the data without relying on a centralized index server. In the latter case, there is no single point of failure in the system.

When the term peer-to-peer was used to describe the Napster network, it implied that the peer protocol was important, but, in reality, the great achievement of Napster was the empowerment of the peers (i.e., the fringes of the network) in association with a central index, which made it fast and efficient to locate available content. The peer protocol was just a common way to achieve this.

Unstructured & Structured P2P Networks

The P2P overlay network consists of all the participating peers as network nodes. There are links between any two nodes that know each other: i.e. if a participating peer knows the location of another peer in the P2P network, then there is a directed edge from the former node to the latter in the overlay network. Based on how the nodes in the overlay network are linked to each other, we can classify the P2P networks as unstructured or structured.

An unstructured P2P network is formed when the overlay links are established arbitrarily. Such networks can be easily constructed as a new peer that wants to join the network can copy existing links of another node and then form its own links over time. In an unstructured P2P network, if a peer wants to find a desired piece of data in the network, the query has to be flooded through the network in order to find as many peers as possible that share the data. The main disadvantage with such networks is that the queries may not always be resolved. Popular content is likely to be available at several peers and any peer searching for it is likely to find the same thing, but if a peer is looking for rare data shared by only a few other peers, then it is highly unlikely that search will be successful. Since there is no correlation between a peer and the content managed by it, there is no guarantee that flooding will find a peer that has the desired data. Flooding also causes a high amount of signalling traffic in the network and hence such networks typically have very poor search efficiency. Most of the popular P2P networks such as Gnutella and FastTrack are unstructured.

Structured P2P networks employ a globally consistent protocol to ensure that any node can efficiently route a search to some peer that has the desired file, even if the file is extremely rare. Such a guarantee necessitates a more structured pattern of overlay links. By far the most common type of structured P2P network is the distributed hash table (DHT), in which a variant of consistent hashing is used to assign ownership of each file to a particular peer, in a way analogous to a traditional hash table's assignment of each key to a particular array slot. Some well known DHTs are Chord, Pastry, Tapestry, CAN, and Tulip.

Legal Controversy

Under US law "the Betamax decision" (Sony Corp. of America v. Universal City Studios, Inc.), case holds that copying "technologies" are not inherently illegal, if substantial non-infringing use can be made of them. This decision, predating the widespread use of the Internet applies to most data networks, including peer-to-peer networks, since distribution of correctly licensed files can be performed. These non-infringing uses include sending open source software, public domain files and out of copyright works. Other jurisdictions tend to view the situation in somewhat similar ways.

In practice, many, often most, of the files shared on peer-to-peer networks are copies of copyrighted popular music and movies. Sharing of these copies among strangers is illegal in most jurisdictions. This has led many observers, including most media companies and some peer-to-peer critics, to conclude that the networks themselves pose grave threats to the established distribution model. The research that attempts to measure actual monetary loss has been somewhat equivocal. Whilst on paper the existence of these networks results in large losses, the actual income does not seem to have changed much since these networks started up. Whether the threat is real or not, both the RIAA and the MPAA now spend large amounts of money attempting to lobby lawmakers for the creation of new laws, and some copyright owners pay companies to help legally challenge users engaging in illegal sharing of their material.

In spite of the Betamax decision, peer-to-peer networks themselves have been targeted by the representatives of those artists and organizations who license their creative works, including industry trade organizations such as the RIAA and MPAA as a potential threat. The Napster service was shut down by an RIAA lawsuit.

In A&M Records v. Napster, 239 F.3d 1004 (9th Cir. 2001), the court found that Napster was both vicariously and contributorily liable for the copyright infringement its users were engaged in. Vicarious liability in these types of cases extends to a provider who financially benefits from the infringement committed by its users while having the ability to police that infringement but has failed to do so. The court found ample evidence that Napster's future revenue is directly dependent upon "increases in userbase." It also found that Napster could have done more than it claimed with regards to restricting users from sharing copyrighted material. Later on in the opinion, the famous peer-to-peer provider was also found contributorily liable. It knew of the infringing use that Napster could and did have, in a medium where the software provided essential access. The RIAA could have sued individual users at the time for violating federal law, but thought it more prudent to shut down the means by which those users shared music.

Napster's use of a central server distinguishes it, on its facts, from the next generation peer-to-peer technology, in which the communication of files is truly "peer to peer." Therefore, additional litigation was needed to determine the legality of their uses.

In MGM v. Grokster, the U.S. Supreme Court reversed a decision of the Ninth Circuit Court of Appeals which had granted a summary judgment of dismissal, and held that there were factual issues concerning whether the defendant p2p software providers had, or had not, encouraged their users to infringe copyrights. If they had done so, they could be held liable for secondary copyright infringement.

The main point of the Grokster holding was about "inducement." "The classic case of direct evidence of unlawful purpose occurs when one induces commission of infringement by another, or "entic[es] or persuad[es] another" to infringe, Black's Law Dictionary 790 (8th ed. 2004), as by advertising. Thus at common law a copyright or patent defendant who "not only expected but invoked [infringing use] by advertisement" was liable for infringement …" The court noted that simply providing the material was not enough. While the decision did not find liability, it simply laid out the groundwork for how the law should be interpreted.

A little over a year later, the RIAA initiated the first major post-Grokster case, Arista v. Limewire, in Manhattan federal court. Lime Wire has counterclaimed in that suit, charging the major record companies with antitrust violations and other misconduct."Lime Wire Sues RIAA for Antitrust Violations"

Shortly thereafter, the lower court judge in Grokster found one of the defendants, Streamcast, the maker of Morpheus, to be liable under the standards enunciated by the Supreme Court. "Streamcast Held Liable for Copyright Infringement in MGM v. Grokster, Round 2" ("Recording Industry vs. The People")

The foregoing cases dealt with 'secondary liability' under the Copyright Act, i.e. whether and when the p2p network software providers can be held liable for 'primary' infringement by their customers. Meanwhile, the underlying question of what uses those customers can make of p2p software without committing a primary infringement is a matter just beginning to be explored, most notably in litigations brought by the Motion Picture Association of America (MPAA) and the Recording Industry Association of America (RIAA) against p2p customers.

In Elektra v. Barker, an RIAA case against Tenise Barker, a Bronx nursing student, Ms. Barker moved to dismiss the complaint, contending, among other things, that the RIAA's allegation of "making available" did not state any known claim under the Copyright Act. The RIAA countered with the argument that "making available" is a copyright infringement, even though the language does not appear in the Act. Thereafter, several amicus curiae were permitted to file briefs in the case, including the MPAA, which agreed with the RIAA's argument, and the Electronic Frontier Foundation (EFF), the U.S. Internet Industry Association (USIIA), and the Computer & Communications Industry Association (CCIA), which agreed with Ms. Barker. The US Department of Justice submitted a brief refuting one of the arguments made by EFF, but did not take any position on the RIAA's "making available" argument. The Elektra v. Barker case was argued before Judge Kenneth M. Karas in Manhattan federal court on January 26, 2007, and as of this writing is awaiting decision.

Anonymous peer-to-peer networks allow for distribution of material - legal or not - with little or no legal accountability across a wide variety of jurisdictions. Many profess that this will lead to greater or easier trading of illegal material and even (as some suggest) facilitate terrorism, and call for its regulation on those grounds [10]. Others counter that the presumption of innocence must apply, and that non peer-to-peer technologies like e-mail (for which there are also anonymizing services), have similar capabilities. Further, the potential for illegal uses should not prevent the technology from being used for legal purposes.

In the European Union (EU), the 2001 EU Copyright directive, which implemented the 1996 WIPO treaty ("World Intellectual Property Organization Copyright Treaty"), prohibits peer-to-peer, claiming it is a violation of the directive. However, not all European member states have implemented the directive in national legislation. Notably, on December 22, 2005, after discussing the EU directive, the French parliament passed two amendments legalizing the exchange of copies on the internet for private use. In a later proceeding, the French government withdrew the article in question and made illegal any p2p client obviously aimed at sharing copyrighted material. The term "obviously" was not defined. The project of law (called DADVSI) has still to be discussed by the French senate and, if the decision differs too much from the Parliament's, it will be debated on second lecture back at the Parliament (Assemblée Nationale).

Interestingly, Canada stands out by authorizing, at least until the projected copyright reform, downloads on peer-to-peer networks under the "private copying" exception.

Computer Science Perspective

Technically, a completely pure peer-to-peer application must implement only peering protocols that do not recognize the concepts of "server" and "client". Such pure peer applications and networks are rare. Most networks and applications described as peer-to-peer actually contain or rely on some non-peer elements, such as DNS. Also, real world applications often use multiple protocols and act as client, server, and peer simultaneously, or over time. Completely decentralized networks of peers have been in use for many years: two examples are Usenet (1979) and FidoNet (1984).

Many P2P systems use stronger peers (super-peers, super-nodes) as servers and client-peers are connected in a star-like fashion to a single super-peer.

Sun added classes to the Java technology to speed the development of peer-to-peer applications quickly in the late 1990s so that developers could build decentralized real time chat applets and applications before Instant Messaging networks were popular. This effort is now being continued with the JXTA project.

Peer-to-peer systems and applications have attracted a great deal of attention from computer science research; some prominent research projects include the Chord project, the PAST storage utility, the P-Grid, a self-organized and emerging overlay network and the CoopNet content distribution system (see below for external links related to these projects).

Application Of P2P Networks Outside Computer Science

- Bioinformatics:Peer-to-peer networks have also begun to attract attention from scientists in other disciplines, especially those that deal with large datasets such as bioinformatics. P2P networks can be used to run large programs designed to carry out tests to identify drug candidates. The first such program was begun in 2001 the Centre for Computational Drug Discovery at Oxford University in cooperation with the National Foundation for Cancer Research. There are now several similar programs running under the auspices of the United Devices Cancer Research Project. On a smaller scale, a self-administered program for computational biologists to run and compare various bioinformatics software is available from Chinook.

- Education and Academic: Due to the fast distribution and large storage space features, many organizations are trying to apply P2P network for educational and academic purposes. For instance, Pennsylvania State University, MIT and Simon Fraser University are carrying on a project called LionShare designed for facilitating file sharing among educational institutions globally.

- Military: The U.S. Department of Defense has already started research topic on P2P network as part of its modern network war. In May, 2003 Dr. Tether. Director of Defense Advanced Research Project Agency has testified that U.S. Military is using P2P network. Due to security reasons, many files are still kept in confidential.

- Business: P2P network has already been used in business areas, but it is still at the beginning line. Currently, Kato et al’s studies indicate over 200 companies with approximately $400 million USD are investing in P2P network. Besides File Sharing, companies are also interested in Distributing Computing, Content Distribution, e-market place, Distributed Search engines, Groupware and Office Automation via P2P network. There are several reasons why companies prefer P2P sometimes such as: Real-time collaboration, a server cannot manage with increasing volume of contents, a process requires strong computing power, a process needs high-speed communications etc. At the same time, P2P is not fully used as it still confronts a lot of security issues.

- TV

- Telecommunication: Nowadays, people are not just satisfied with “can hear a person from another side of the earth”, instead, the demands of clearer voice in real-time are increasing globally. Just like the TV network, there are already cables built. It’s not very likely for companies to change all the cables. Many of them turn to use internet, more specifically, P2P network. For instance, Skype, one of the most widely used phone software is using P2P technology. Furthermore, many research organizations are trying to apply P2P network on cellular network.

Attacks On Peer-To-Peer Networks

Many peer-to-peer networks are under constant attack by people with a variety of motives.

Examples include:

- poisoning attacks (e.g. providing files whose contents are different from the description)
- polluting attacks (e.g. inserting "bad" chunks/packets into an otherwise valid file on the network)
- defection attacks (users or software that make use of the network without contributing resources to it)
- insertion of viruses to carried data (e.g. downloaded or carried files may be infected with viruses or other malware)
- malware in the peer-to-peer network software itself (e.g. distributed software may contain spyware)
- denial of service attacks (attacks that may make the network run very slowly or break completely)
- filtering (network operators may attempt to prevent peer-to-peer network data from being carried)
- identity attacks (e.g. tracking down the users of the network and harassing or legally attacking them)
- spamming (e.g. sending unsolicited information across the network- not necessarily as a denial of service attack)

Most attacks can be defeated or controlled by careful design of the peer-to-peer network and through the use of encryption. P2P network defence is in fact closely related to the "Byzantine Generals Problem". However, almost any network will fail when the majority of the peers are trying to damage it, and many protocols may be rendered impotent by far fewer numbers.

The Byzantine Generals Problem (BGP)

Byzantine fault tolerance is the name given to a sub-field of error tolerance research, inspired by The Byzantine Generals' Problem, which is a generalized version of the Two Generals' Problem.

The object of Byzantine fault tolerance is to be able to defend against a Byzantine failure, in which a component of some system not only behaves erroneously, but also fails to behave consistently when interacting with multiple other components. Correctly functioning components of a Byzantine fault tolerant system will be able to reach the same group decisions regardless of Byzantine faulty components. There are upper bounds on the percentage of traitorous or unreliable components, however.

BGP: Byzantine Failures

A Byzantine failure (or Byzantine fault) is an arbitrary fault that occurs during the execution of an algorithm by a distributed system. It encompasses those faults that are commonly referred to as "crash failures" and "send and omission failures". When a Byzantine failure has occurred, the system may respond in any unpredictable way, unless it is designed to have Byzantine fault tolerance.

These arbitrary failures may be loosely categorized as follows:

- a failure to take another step in the algorithm, also known as a crash failure;
- a failure to correctly execute a step of the algorithm; and
- arbitrary execution of a step other than the one indicated by the algorithm.

For example, if the output of one function is the input of another, then small round-off errors in the first function can produce much larger errors in the second. If the second function were fed into a third, the problem could grow even larger, until the values produced are worthless. Another example is in compiling source code. One minor syntactical error early on in the code can produce large numbers of perceived errors later, as the compiler gets out-of-phase with the lexical and syntactic information in the source program.

Steps are taken by processes, the abstractions that execute the algorithms. A faulty process is one that at some point exhibits one of the above failures. A process that is not faulty is correct.

The Byzantine failure assumption models real-world environments in which computers and networks may behave in unexpected ways due to hardware failures, network congestion and disconnection, as well as malicious attacks. Byzantine failure-tolerant algorithms must cope with such failures and still satisfy the specifications of the problems they are designed to solve. Such algorithms are commonly characterized by their resilience t, the number of faulty processes with which an algorithm can cope.

Many classic agreement problems, such as the Byzantine Generals Problem, have no solution unless $t < n / 3$, where n is the number of processes in the system.

The Two Generals' Problem is a specific case which assumes that processes are reliable but communication between processes is not reliable.

BGP: Origin

Byzantine refers to the Byzantine Generals' Problem, an agreement problem in which generals of the Byzantine Empire's army must decide unanimously whether or not to attack some enemy army. The problem is complicated by the geographic separation of the generals, who must communicate by sending messengers to each other, and by the presence of traitors amongst the generals. These traitors can act arbitrarily in order to achieve the following aims: trick some generals into attacking; force a decision that is not consistent with the generals' desires, e.g. forcing an attack when no general wished to attack; or confusing some generals to the point that they are unable to make up their minds. If the traitors succeed in any of these goals, any resulting attack is doomed, as only a concerted effort can result in victory.

Byzantine fault tolerance can be achieved if the loyal (non-faulty) generals have a unanimous agreement on their strategy. Note that if the source general is correct, all loyal generals must agree upon that value. Otherwise, the choice of strategy agreed upon is irrelevant.

BGP: Solutions

Several solutions were originally described by Lamport, Shostak, and Pease in 1982. They began by noting that the Generals Problem can be reduced to solving a "Commander and Lieutenants" problem where Loyal Lieutenants must all act in unison and that their action must correspond to what the Commander ordered in the case that the Commander is Loyal. Roughly speaking, the Generals vote by treating each others' orders as votes.

- One solution considers scenarios in which messages may be forged, but which will be Byzantine Fault Tolerant as long as the number of traitorous generals does not equal or exceed one third. The impossibility of dealing with one-third or more traitors ultimately reduces to proving that the 1 Commander + 2 Lieutenants problem cannot be solved if the Commander is traitorous. The reason is, if we have three commanders, A, B, and C, and A is the traitor: when A tells B to attack and C to retreat, and B and C sends messages to each other, forwarding A's message, neither B nor C can figure out who is the traitor, since it isn't necessarily A - the other commander could have forged the message purportedly from A. It can be shown that if n is the number of generals in total, and t is the number of traitors in that n, then there are solutions to the problem only when n is greater than or equal to 3 times t + 1.

- A second solution requires unforgeable signatures (in modern computer systems, this may be achieved through public key cryptography), but maintains Byzantine Fault Tolerance in the presence of an arbitrary number of traitorous generals.

- Also presented is a variation on the first two solutions allowing Byzantine Fault Tolerant behavior in some situations where not all generals can communicate directly with each other.

Data & Networking Practises – Reliability & Availability

(Source: EventHelix.com)

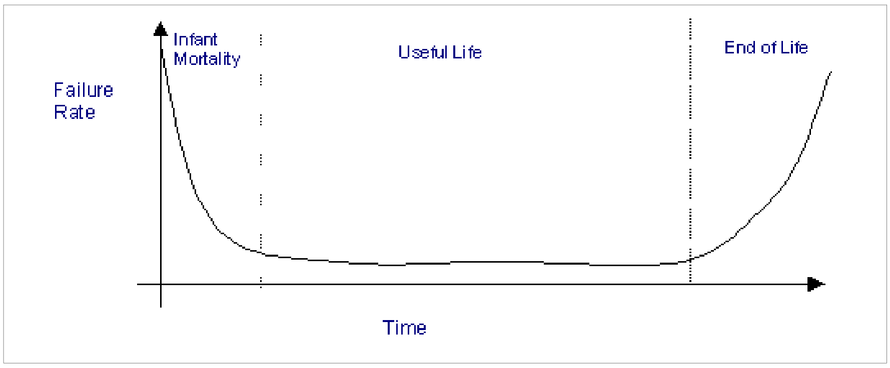
Reliability & Availability Basics

Real-time and embedded systems are now a central part of our lives. Reliable functioning of these systems is of paramount concern to the millions of users that depend on these systems everyday. Unfortunately most embedded systems still fall short of user's expectation of reliability.

In this article we will discuss basic techniques for measuring and improving reliability of computer systems. The following topics are discussed:

Hardware Failures

Hardware failures are typically characterized by a bath tub curve. An example curve is shown below. The chance of a hardware failure is high during the initial life of the module. The failure rate during the rated useful life of the product is fairly low. Once the end of the life is reached, failure rate of modules increases again.



Hardware failures during a products life can be attributed to the following causes:

Design failures	This class of failures take place due to inherent design flaws in the system. In a well designed system this class of failures should make a very small contribution to the total number of failures.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

Infant Mortality	This class of failures cause newly manufactured hardware to fail. This type of failures can be attributed to manufacturing problems like poor soldering, leaking capacitor etc. These failures should not be present in systems leaving the factory as these faults will show up in factory system burn in tests.
Random Failures	Random failures can occur during the entire life of a hardware module. These failures can lead to system failures. Redundancy is provided to recover from this class of failures.
Wear Out	Once a hardware module has reached the end of its useful life, degradation of component characteristics will cause hardware modules to fail. This type of faults can be weeded out by preventive maintenance and routing of hardware.

Software Failures

Software failures can be characterized by keeping track of software defect density in the system. This number can be obtained by keeping track of historical software defect history. Defect density will depend on the following factors:

- Software process used to develop the design and code (use of peer level design/code reviews, unit testing)
- Complexity of the software
- Size of the software
- Experience of the team developing the software
- Percentage of code reused from a previous stable project
- Rigor and depth of testing before product is shipped.

Defect density is typically measured in number of defects per thousand lines of code (defects/ KLOC).

Software Fault Tolerance

Most Real-time systems focus on hardware fault tolerance. Software fault tolerance is often overlooked. This is really surprising because hardware components have much higher reliability than the software that runs over them. Most system designers go to great lengths to limit the impact of a hardware failure on system performance. However they pay little attention to the systems behaviour when a software module fails.

In this article we will be covering several techniques that can be used to limit the impact of software faults (read bugs) on system performance. The main idea here is to contain the damage caused by software faults. Software fault tolerance is not a license to ship the system with bugs. The real objective is to improve system performance and availability in cases when the system encounters a software or hardware fault.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

- Timeouts
- Audits
- Exception Handling
- Task Rollback
- Incremental Reboot
- Voting

Timeouts

Most Real-time systems use timers to keep track of feature execution. A timeout generally signals that some entity involved in the feature has misbehaved and a corrective action is required. The corrective action could be of two forms:

- Retry: When the application times out for a response, it can retry the message interaction. You might argue that we do not need to implement application level retries as lower level protocols will automatically recover from message loss. Keep in mind that message loss recovery is not the only objective of implementing retries. Retries help in recovering from software faults too. Consider a scenario where a message sent to a task is not processed because of a task restart or processor reboot. An application level retry will recover from this condition.

- Abort: In this case timeout for a response leads to aborting of the feature. This might seem too drastic, but in reality aborting a feature might be the simplest and safest solution in recovering from the errors. The feature might be retried by the user invoking the feature. Consider a case where a call has to be cleared because the task originating the call did not receive a response in time. If this condition can happen only in rare scenarios, the simplest action on timeout might be to clear the call. The user would retry the call.

The choice between retrying or aborting on timeouts is based on several factors. Consider all these factors before you decide either way:

- If the feature being executed is fairly important for system stability, it might be better to retry. For example, a system start-up feature should not be aborted on one timeout.
- If the lower layer protocol is not robust, retry might be a good option. For example, message interactions using an inherently unreliable protocol like slotted aloha should always be retried.
- Complexity of implementation should also be considered before retrying a message interaction. Aborting a feature is a simpler option. More often than not system designers just default to retrying without even considering the abort option. Keep in mind that retry implementation complicates the code and state machine design.
- If the entity invoking this feature will retry the feature, the simplest action might be abort the feature and wait for an external retry.
- Retrying every message in the system will lower system performance because of frequent timer start and stop operations. In many cases, performance can be improved by just running a single timer for the complete feature execution. On timeout the feature can simply be aborted.
- For most external interactions, the designer might have no choice. As the timeouts and retry actions are generally specified by the external protocols.

- Many times the two techniques are used together. The task retries a message certain number of times. If no response is received after exhausting this limit, the feature might be aborted.

Audits

Most Real-time systems comprise of software running across multiple processors. This implies that data is also distributed. The distributed data may get inconsistent in Real-time due to reasons like:

- independent processor reboot
- software bugs
- race conditions
- hardware failures
- protocol failures

The system must behave reliably under all these conditions. A simple strategy to overcome data inconsistency is to implement audits. Audit is a program that checks the consistency of data structures across processors by performing predefined checks.
Audit Procedure

1. System may trigger audits due to several reasons:
    - periodically
    - failure of certain features
    - processor reboots
    - processor switchovers
    - certain cases of resource congestion
2. Audits perform checks on data and look for data inconsistencies between processors.
3. Since audits have to run on live systems, they need to filter out conditions where the data inconsistency is caused by transient data updates. On data inconsistency detection, audits perform multiple checks to confirm inconsistency. An inconsistency is considered valid if and only if it is detected on every iteration of the check.
4. When inconsistency is confirmed, audits may perform data structure cleanups across processors.
5. At times audits may not directly cleanup inconsistencies; they may trigger appropriate feature aborts etc.

An Example

Let’s consider the Xenon Switching System. If the call occupancy on the system is much less than the maximum that could be handled and still calls are failing due to lack of space-slot resources, call processing subsystem will detect this condition and will trigger space-slot audit. The audit will run on the XEN and CAS processors crosscheck if a space-slot that is busy at CAS actually has a corresponding call at XEN. If no active call is found on XEN for a space-slot, the audit will recheck the condition after a small delay for several times. If the inconsistency holds on every attempt, the space-slot resource is marked free at CAS. The audit performs several rechecks to eliminate the scenario in which the space-slot release message may be in transit.

Exception Handling

Whenever a task receives a message, it performs a series of defensive checks before processing it. The defensive checks should verify the consistency of the message as well as the internal state of the task. Exception handler should be invoked on defensive check failure.

Depending on the severity, exception handler can take any of the following actions:

- Log a trace for developer post processing.
- Increment a leaky-bucket counter for the error condition.
- Trigger appropriate audit.
- Trigger a task rollback.
- Trigger processor reboot.

Leaky Bucket Counter

Leaky-bucket counters are used to detect a flurry of error conditions. To ignore rare error conditions they are periodically leaked i.e. decremented. If these counters reach a certain threshold, appropriate exception handling is triggered. Note that the threshold will never be crossed by rare happening of the associated error condition. However, if the error condition occurs rapidly, the counter will overflow i.e. cross the threshold.

Task Rollback

In a complex Real-time system, a software bug in one task leading to processor reboot may not be acceptable. A better option in such cases is to isolate the erroneous task and handle the failure at the task level. The task in turn may decide to rollback i.e. start operation from a known or previously saved state. In other cases, it may not be expensive to forget the context by just deleting the offending task and informing other associated tasks.

For example, if the Space Slot Manager on the CAS card encounters a exception condition leading to task rollback, it might resume operation by recovering the space slot allocation status from the connection memory. On the other hand, exception in a call task might just be handled by clearing the call task and releasing all the resources assigned to this task.

Task rollback may be triggered by any of the following events:

- Hardware exception conditions like divide by zero, illegal address access (bus error)
- Defensive check leaky-bucket counter overflows.
- Audit detected inconsistency to be resolved by task rollback.

Incremental Reboot

Software processor reboots can be time consuming, leading to unacceptable amount of downtime. To reduce the system reboot time, complex Real-time systems often implement incremental system initialisation procedures. For example, a typical Real-time system may implement three levels of system reboot:

- Level 1 Reboot : Operating system reboot
- Level 2 Reboot : Operating system reboot along with configuration data download
- Level 3 Reboot: Code reload followed by operating system reboot along with configuration data download.

Incremental Reboot Procedure

1. A defensive check leaky-bucket counter overflow will typically lead to rollback of the offending task.
2. In most cases task rollback will fix the problem. However, in some cases, the problem may not be fixed leading to subsequent rollbacks too soon. This will cause the task level rollback counter to overflow, leading to a Level 1 Reboot.
3. Most of the times, Level 1 Reboot will fix the problem. But in some cases, the processor may continue to hit Level 1 Reboots repeatedly. This will cause the Level 1 Reboot counter to overflow, leading to a Level 2 Reboot.
4. Majority of the times, Level 2 Reboot is able to fix the problem. If it is unable to fix the problem, the processor will repeatedly hit Level 2 Reboots, causing the Level 2 Reboot counter to overflow leading to Level 3 Reboot.

Voting

This is a technique that is used in mission critical systems where software failure may lead to loss of human life .e.g. aircraft navigation software. Here, the Real-time system software is developed by at least three distinct teams. All the teams develop the software independently. And, in a live system, all the three implementations are run simultaneously. All the inputs are fed to the three versions of software and their outputs are voted to determine the actual system response. In such systems, a bug in one of the three modules will get voted out by the other two versions.

Reliability Parameters

We distinguish between three types;

- MTBF
- FITS
- MTTR

MTBF

Mean Time Between Failures (MTBF), as the name suggests, is the average time between failures of hardware modules. It is the average time a manufacturer estimates before a failure occurs in a hardware module.

MTBF for hardware modules can be obtained from the vendor for off-the-shelf hardware modules. MTBF for in-house developed hardware modules is calculated by the hardware team developing the board.

MTBF for software can be determined by simply multiplying the defect rate with KLOCs executed per second.

FITS

FITS is a more intuitive way of representing MTBF. FITS is nothing but the total number of failures of the module in a billion hours (i.e. 1000,000,000 hours).

MTTR

Mean Time To Repair (MTTR) is the time taken to repair a failed hardware module. In an operational system, repair generally means replacing the hardware module. Thus hardware

MTTR could be viewed as mean time to replace a failed hardware module. It should be a goal of system designers to allow for a high MTTR value and still achieve the system reliability goals. You can see from the table below that a low MTTR requirement means high operational cost for the system.

Table: Estimating The Hardware MTTR

Where are hardware spares kept?	How is site manned?	Estimated MTTR
Onsite	24 hours a day	30 minutes
Onsite	Operator is on call 24 hours a day	2 hours
Onsite	Regular working hours on week days as well as weekends and holidays	14 hours
Onsite	Regular working hours on week days only	3 days
Offsite. Shipped by courier when fault condition is encountered.	Operator paged by system when a fault is detected.	1 week
Offsite. Maintained in an operator controlled warehouse	System is remotely located. Operator needs to be flown in to replace the hardware.	2 week

MTTR for a software module can be computed as the time taken to reboot after a software fault is detected. Thus software MTTR could be viewed as the mean time to reboot after a software fault has been detected. The goal of system designers should be to keep the software MTTR as low as possible. MTTR for software depends on several factors:

- Software fault tolerance techniques used
- OS selected (does the OS allow independent application reboot?)
- Code image downloading techniques

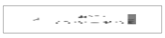
Table: Estimating The Hardware MTTR

Software fault recovery mechanism	Software reboot mechanism on fault detection	Estimated MTTR
Software failure is detected by watchdog and/or health messages	Processor automatically reboots from a ROM resident image.	30 seconds
Software failure is	Processor	30 seconds

Software fault recovery mechanism	Software reboot mechanism on fault detection	Estimated MTTR
detected by watchdog and/or health messages	automatically restarts the offending tasks, without needing an operating system reboot	
Software failure is detected by watchdog and/or health messages	Processor automatically reboots and the operating system reboots from disk image and restarts applications	3 minutes
Software failure is detected by watchdog and/or health messages	Processor automatically reboots and the operating system and application images have to be download from another machine	10 minutes
Software failure detection is not supported.	Manually operator reboot is required.	30 minutes to 2 weeks (software MTTR is same as hardware MTTR)

Availability

Availability of the module is the percentage of time when system is operational. Availability of a hardware/software module can be obtained by the formula given below.



Availability is typically specified in nines notation. For example 3-nines availability corresponds to 99.9% availability. A 5-nines availability corresponds to 99.999% availability.

Downtime

Downtime per year is a more intuitive way of understanding the availability. The table below compares the availability and the corresponding downtime.

Availability	Downtime
90% (1-nine)	36.5 days/year
99% (2-nines)	3.65 days/year
99.9% (3-nines)	8.76 hours/year
99.99% (4-nines)	52 minutes/year
99.999% (5-nines)	5 minutes/year

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

99.9999% (6-nines)	31 seconds/year!

Data & Networking Practises – BitTorrent

High-Level

BitTorrent is a peer-to-peer file sharing (P2P) communications protocol. BitTorrent is a method of distributing large amounts of data widely without the original distributor incurring the entire costs of hardware, hosting, and bandwidth resources. Instead, when data is distributed using the BitTorrent protocol, each recipient supplies pieces of the data to newer recipients, reducing the cost and burden on any given individual source, providing redundancy against system problems, and reducing dependence on the original distributor.

The protocol is the brainchild of programmer Bram Cohen, who designed it in April 2001 and released a first implementation on 2 July 2001. It is now maintained by Cohen's company BitTorrent, Inc.

Usage of the protocol accounts for significant traffic on the Internet, but the precise amount has proven difficult to measure.

There are numerous compatible BitTorrent clients, written in a variety of programming languages, and running on a variety of computing platforms.

Operation

A BitTorrent client is any program that implements the BitTorrent protocol. Each client is capable of preparing, requesting, and transmitting any type of computer file over a network, using the protocol. A peer is any computer running an instance of a client.

To share a file or group of files, a peer first creates a small file called a "torrent" (e.g. MyFile.torrent). This file contains metadata about the files to be shared and about the tracker, the computer that coordinates the file distribution. Peers that want to download the file first obtain a torrent file for it, and connect to the specified tracker, which tells them from which other peers to download the pieces of the file.

Though both ultimately transfer files over a network, a BitTorrent download differs from a classic full-file HTTP request in several fundamental ways:

* BitTorrent makes many small data requests over different TCP sockets, while web-browsers typically make a single HTTP GET request over a single TCP socket.
* BitTorrent downloads in a random or in a "rarest-first" approach that ensures high availability, while HTTP downloads in a sequential manner.

Taken together, these differences allow BitTorrent to achieve much lower cost, much higher redundancy, and much greater resistance to abuse or to "flash crowds" than a regular HTTP server. However, this protection comes at a cost: downloads can take time to rise to full speed because it may take time for enough peer connections to be established, and it takes time for a node to receive sufficient data to become an effective uploader. As such, a typical BitTorrent download will gradually rise to very high speeds, and then slowly fall back down toward the end of the download. This contrasts with an HTTP server that, while more vulnerable to overload and abuse, rises to full speed very quickly and maintains this speed throughout.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited.*io

In general, BitTorrent's non-contiguous download methods have prevented it from supporting "progressive downloads" or "streaming playback". But recent comments by Bram Cohen suggest that streaming torrent downloads will soon be commonplace and ad supported streaming appears to be the result of those comments.

Creating & Publishing Torrents

The peer distributing a data file treats the file as a number of identically-sized pieces, typically between 64 kB and 4 MB each. The peer creates a checksum for each piece, using the SHA1 hashing algorithm, and records it in the torrent file. Pieces with sizes greater than 512 kB will reduce the size of a torrent file for a very large payload, but is claimed to reduce the efficiency of the protocol. When another peer later receives a particular piece, the checksum of the piece is compared to the recorded checksum to test that the piece is error-free. Peers that provide a complete file are called seeders, and the peer providing the initial copy is called the initial seeder.

The exact information contained in the torrent file depends on the version of the BitTorrent protocol. By convention, the name of a torrent file has the suffix .torrent. Torrent files have an "announce" section, which specifies the URL of the tracker, and an "info" section, containing (suggested) names for the files, their lengths, the piece length used, and a SHA-1 hash code for each piece, all of which is used by clients to verify the integrity of the data they receive.

Completed torrent files are typically published on websites or elsewhere, and registered with a tracker. The tracker maintains lists of the clients currently participating in the torrent. Alternatively, in a trackerless system (decentralized tracking) every peer acts as a tracker. This is implemented by the BitTorrent, µTorrent, BitComet, KTorrent and Deluge clients through the distributed hash table (DHT) method. Azureus also supports a trackerless method that is incompatible (as of April 2007) with the DHT offered by all other supporting clients.

In November 2006, BitTorrent Inc. introduced its "Publish Torrent" service, which creates and hosts a torrent file (seeded from an existing web-hosted media file) and tracks the downloads. The service (http://www.bittorrent.com/publish) requires a client that supports web-seeding (currently the official client, Azureus, µTorrent and anything based on Libtorrent).

Downloading Torrents & Sharing Files

Users browse the web to find a torrent of interest, download it, and open it with a BitTorrent client. The client connects to the tracker(s) specified in the torrent file, from which it receives a list of peers currently transferring pieces of the file(s) specified in the torrent. The client connects to those peers to obtain the various pieces. Such a group of peers connected to each other to share a torrent is called a swarm. If the swarm contains only the initial seeder, the client connects directly to it and begins to request pieces. As peers enter the swarm, they begin to trade pieces with one another, instead of downloading directly from the seeder.

Clients incorporate mechanisms to optimize their download and upload rates; for example they download pieces in a random order to increase the opportunity to exchange data, which is only possible if two peers have different pieces of the file.

The effectiveness of this data exchange depends largely on the policies that clients use to determine to whom to send data. Clients may prefer to send data to peers that send data back to them (a tit for tat scheme), which encourages fair trading. But strict policies often result in suboptimal situations; e.g., when newly joined peers are unable to receive any data because they don't have any pieces yet to trade themselves or when two peers with a good connection between them do not exchange data simply because neither of them wants to take the initiative. To counter these effects, the official BitTorrent client program uses a mechanism called “optimistic unchoking,” where the client reserves a portion of its available bandwidth for sending pieces to random peers (not necessarily known-good partners, so called preferred peers), in hopes of discovering even better partners and to ensure that newcomers get a chance to join the swarm.

Adoption

A growing number of individuals and organizations are using BitTorrent to distribute their own or licensed material. Independent adopters report that without using BitTorrent technology, and its dramatically reduced demands on networking hardware and bandwidth, they could not afford to distribute their files.

Film, Video & Music

- BitTorrent Inc. has amassed a number of licenses from Hollywood studios for distributing popular content at the company's website.
- Sub Pop Records releases tracks and videos via BitTorrent Inc. to distribute its 1000+ albums. The band Ween uses the website Browntracker.net to distribute free audio and video recordings of live shows. Furthermore, Babyshambles and The Libertines (both bands associated with Pete Doherty) have extensively used torrents to distribute hundreds of demos and live videos.
- Podcasting software is starting to integrate BitTorrent to help podcasters deal with the download demands of their MP3 "radio" programs. Specifically, Juice and Miro (formerly known as Democracy Player) support automatic processing of .torrent files from RSS feeds. Similarly, some BitTorrent clients, such as µTorrent, are able to process web feeds and automatically download content found within them.

Broadcasters

- In 2008 CBC became the first public broadcaster in North America to make a full show (Canada's Next Great Prime Minister) available for download using BitTorrent.
- The Norwegian Broadcasting Corporation (NRK) have since March 2008 experimented with bittorrent distribution from this site. Only selected material in which NRK owns all royalties are published. Responses have been very positive, and NRK is planning to offer more content.

Personal Material

- The Amazon S3 "Simple Storage Service" is a scalable Internet-based storage service with a simple web service interface, equipped with built-in BitTorrent support.
- Blog Torrent offers a simplified BitTorrent tracker to enable bloggers and non-technical users to host a tracker on their site. Blog Torrent also allows visitors to download a "stub" loader, which acts as a BitTorrent client to download the desired file, allowing users without BitTorrent software to use the protocol. This is similar to the concept of a self-extracting archive.

Software

- Many major open source and free software projects encourage BitTorrent as well as conventional downloads of their products to increase availability and reduce load on their own servers.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

Games

- Blizzard's World of Warcraft video game utilizes the BitTorrent protocol to send game updates to clients.
- The game GunZ The Duel has a built-in BitTorrent client.
- Metal Gear Online uses BitTorrent has a built-in BitTorrent update downloader and patcher, which is recommended by KONAMI to be used for higher download speeds, when the update menu appears and the BitTorrent button is highlighted.

Network Impact

CableLabs, the research organization of the North American cable industry, estimates that BitTorrent represents 18% of all broadband traffic. In 2004, CacheLogic put that number at roughly 35% of all traffic on the Internet. The discrepancies in these numbers are caused by differences in the methodology used to measure P2P traffic on the Internet. Routers that use NAT, Network Address Translation, must maintain tables of source and destination IP addresses and ports. Typical home routers are limited to about 2000 table entries while some more expensive routers have larger table capacities. BitTorrent frequently contacts 300-500 servers per second rapidly filling the NAT tables. This is a common cause of home routers locking up.

Indexing

The BitTorrent protocol provides no way to index torrent files. As a result, a comparatively small number of websites have hosted the large majority of torrents linking to (possibly) copyrighted material, rendering those sites especially vulnerable to lawsuits. Several types of websites support the discovery and distribution of data on the BitTorrent network.

Public tracker sites such as The Pirate Bay allow users to search in and download from their collection of .torrent files; they also run BitTorrent trackers for those files. Users can typically also upload .torrent files for content they wish to distribute.

Private tracker sites such as Demonoid operate like public ones except that they restrict access to registered users and keep track of the amount of data each user uploads and downloads, in an attempt to reduce leeching.

There are specialized tracker sites such as FlixFlux for films, bitme for educational content, PureTnA for pornographic content, and tv torrents for television series. Often these will also be private.

Search engines allow the discovery of .torrent files that are hosted and tracked on other sites; examples include Mininova, Monova, Btjunkie, Torrentz and isoHunt. These sites allow the user to ask for content meeting specific criteria (such as containing a given word or phrase) and retrieve a list of links to .torrent files matching those criteria. This list is often sorted with respect to relevance or number of seeders. Bram Cohen launched a BitTorrent search engine on http://search.bittorrent.com that commingles licensed content with search results. Metasearch engines allow to search several BitTorrent indices and search engines at once.

Technologies Built On BitTorrent

The BitTorrent protocol is still under development and therefore may still acquire new features and other enhancements such as improved efficiency.

Distributed Trackers

In June 2005, BitTorrent, Inc. released version 4.2.0 of the Mainline BitTorrent client. This release supported "trackerless" torrents, featuring a DHT implementation which allowed the client to use torrents that do not have a working BitTorrent tracker. Current versions of the official BitTorrent client, µTorrent, BitComet, and BitSpirit all share a compatible DHT implementation that is based on Kademlia. Azureus uses its own incompatible DHT system called the "distributed database", but a plugin is available which allows use of the mainline DHT.

Another idea that has surfaced recently in Azureus is that of virtual torrents. This idea is based on the distributed tracker approach and is used to describe some web resource. Currently, it is used for instant messaging. It is implemented using a special messaging protocol and requires an appropriate plugin. Anatomic P2P is another approach, which uses a decentralized network of nodes that route traffic to dynamic trackers.

Most BitTorrent clients also use Peer exchange (PEX) to gather peers in addition to trackers and DHT. Peer exchange checks with known peers to see if they know of any other peers. With the 3.0.5.0 release of Azureus, all major BitTorrent clients now have compatible peer exchange.

Content Delivery

Web seeding was implemented in 2006. The advantage of this feature is that a site may distribute a torrent for a particular file or batch of files and make those files available for download from that same web server; this can simplify seeding and load balancing greatly once support for this feature is implemented in the various BitTorrent clients. In theory, this would make using BitTorrent almost as easy for a web publisher as simply creating a direct download while allowing some of the upload bandwidth demands to be placed upon the downloaders (who normally use only a very small portion of their upload bandwidth capacity). This feature was created by John "TheSHAD0W" Hoffman, who created BitTornado. From version 5.0 onward the Mainline BitTorrent client also supports web seeds and the BitTorrent web site has a simple publishing tool that creates web seeded torrents. µTorrent added support for web seeds in version 1.7. The latest version of the popular download manager GetRight supports downloading a file from both HTTP/FTP protocols and using BitTorrent.

Broadcatching combines RSS with the BitTorrent protocol to create a content delivery system, further simplifying and automating content distribution. Steve Gillmor explained the concept in a column for Ziff-Davis in December, 2003. The discussion spread quickly among bloggers (Techdirt, Ernest Miller, Chris Pirillo, etc.). In an article entitled Broadcatching with BitTorrent, Scott Raymond explained:

> *"I want RSS feeds of BitTorrent files. A script would periodically check the feed for new items, and use them to start the download. Then, I could find a trusted publisher of an Alias RSS feed, and 'subscribe' to all new episodes of the show, which would then start downloading automatically — like the 'season pass' feature of the TiVo."*

The RSS feed will track the content, while BitTorrent ensures content integrity with cryptographic hashing of all data, so subscribers to a feed receive uncorrupted content.

One of the first software clients (free and open source) for broadcatching is Miro. Other free software clients such as PenguinTV and KatchTV are also now supporting broadcatching.

The BitTorrent web-service MoveDigital has the ability to make torrents available to any web application capable of parsing XML through its standard Representational State Transfer (REST) based interface. Additionally, Torrenthut is developing a similar torrent API that will provide the same features, as well as further intuition to help bring the torrent community to Web 2.0 standards. Alongside this release is a first PHP application built using the API called PEP, which will parse any Really Simple Syndication (RSS 2.0) feed and automatically create and seed a torrent for each enclosure found in that feed.

Encryption

Some ISPs throttle (reduce) BitTorrent traffic of their customers because it makes up a large proportion of total traffic and the ISPs don't want to spend money purchasing extra capacity.

Protocol header encrypt (PHE) and Message stream encryption/Protocol encryption (MSE/PE) are features of some BitTorrent clients that attempt to make BitTorrent hard to detect and throttle. At the moment Azureus, Bitcomet, KTorrent, Transmission, Deluge, µTorrent,MooPolice, Halite, rtorrent and the latest official BitTorrent client (v6) support MSE/PE encryption.

In September 2006 it was reported that some software could detect and throttle BitTorrent traffic masquerading as HTTP traffic.

Reports in August 2007 indicated that Comcast was preventing BitTorrent seeding by monitoring and interfering with the communication between peers. Protection against these efforts is provided by proxying the client-tracker traffic through the Tor anonymity network or, via an encrypted tunnel to a point outside of the Comcast network.

In general, although encryption can make it difficult to determine what is being shared, BitTorrent is generally vulnerable to traffic analysis. Thus even with MSE/PE, it may be possible for an ISP to recognize BitTorrent and also to determine that a system is no longer downloading, only uploading, information and terminate its connection by injecting TCP RST (reset flag) packets.

Multitracker

Another unofficial feature is an extension to the BitTorrent metadata format proposed by John Hoffman and implemented by several indexing websites. It allows the use of multiple trackers per file, so if one tracker fails, others can continue supporting file transfer. It is implemented in several clients, such as BitComet, BitTornado, KTorrent and µTorrent. Trackers are placed in groups, or tiers, with a tracker randomly chosen from the top tier and tried, moving to the next tier if all the trackers in the top tier fail.

Torrents with multiple trackers can decrease the time it takes to download a file, but also has a few consequences:

- Users have to contact more trackers, leading to more overhead-traffic.
- Torrents from closed trackers suddenly become downloadable by non-members, as they can connect to a seed via an open tracker.

Limitations & Security Vulnerabilities

BitTorrent does not offer its users anonymity. It is possible to obtain the IP addresses of all current, and possibly previous, participants in a swarm from the tracker. This may expose users with insecure systems to attacks.

Another drawback is that BitTorrent file sharers, compared to users of client/server technology, often have little incentive to become seeders after they finish downloading. The result of this is that torrent swarms gradually die out, meaning a lower possibility of obtaining older torrents. Some BitTorrent websites have attempted to address this by recording each user's download and upload ratio for all or just the user to see, as well as the provision of access to newer torrent files to people with better ratios. Also, users who have low upload ratios may see slower download speeds until they upload more. This prevents (statistical) leeching, since after a while they become unable to download much faster than 1-10 kB/s on a high-speed connection. Some trackers exempt dial-up users from this policy, because they cannot upload faster than 1-3 kB/s.

There are "cheating" clients like BitThief which claim to be able to download without uploading, and because of this can sometimes download faster than regular clients. Such exploitation negatively affects the cooperative nature of the BitTorrent protocol.

BitTorrent is best suited to continuously connected broadband environments, since dial-up users find it less efficient due to frequent disconnects and slow download rates.

Implementations

Because of the open nature of the protocol, many clients have been developed that support numerous platforms and written using various programming languages. The official client is also named BitTorrent.

Some clients, like Torrentflux, can be run straight from a server, allowing hosting companies to offer speeds unavailable to most users. Sites such as Torrent2FTP offer services to download torrents and then make them available to the customer on a FTP server.

Opera Software now incorporates BitTorrent downloads through its popular browser software, as does Wyzo.

An increasing number of hardware devices are being made to support BitTorrent. These include routers and NAS devices, as well as anything capable of running OpenWrt (routers) or Openslug (NAS).

Development

An as-yet unimplemented unofficial feature is Similarity Enhanced Transfer (SET), a technique for improving the speed at which peer-to-peer file sharing and content distribution systems can share data. SET, proposed by researchers Pucha, Andersen, and Kaminsky, works by spotting chunks of identical data in files that are an exact or near match to the one needed and transferring these data to the client if the 'exact' data are not present. Their experiments suggested that SET will help greatly with less popular files, but not as much for popular data, where many peers are already downloading it. Andersen believes that this technique could be immediately used by developers with the BitTorrent file sharing system.

Data & Networking Practises – Servents

What Is It?

In general a servent is a peer-to-peer network node, which has the functionalities of both a server and a client. This is a portmanteau derived from the terms server and client, and is a play on the word "servant". The setup is designed so that each node can upload, download, and usually also route network information, allowing for the creation and maintenance of ad-hoc networks.

The term originated from the Gnutella lexicon, with Gnutella being the first widespread decentralized peer-to-peer network.

Data & Networking Practises – Database

What Is It?

In computing, a database can be defined as a structured collection of records or data that is stored in a computer so that a program can consult it to answer queries. The records retrieved in answer to queries become information that can be used to make decisions. The computer program used to manage and query a database is known as a database management system (DBMS). The properties and design of database systems are included in the study of information science.

The term "database" originated within the computing discipline. Although its meaning has been broadened by popular use, even to include non-electronic databases, this article is about computer databases. Database-like records have been in existence since well before the Industrial Revolution in the form of ledgers, sales receipts and other business-related collections of data.

The central concept of a database is that of a collection of records, or pieces of Information. Typically, for a given database, there is a structural description of the type of facts held in that database: this description is known as a schema. The schema describes the objects that are represented in the database, and the relationships among them. There are a number of different ways of organizing a schema, that is, of modeling the database structure: these are known as database models (or data models). The model in most common use today is the relational model, which in layman's terms represents all information in the form of multiple related tables each consisting of rows and columns (the true definition uses mathematical terminology). This model represents relationships by the use of values common to more than one table. Other models such as the hierarchical model and the network model use a more explicit representation of relationships.

The term database refers to the collection of related records, and the software should be referred to as the database management system or DBMS. When the context is unambiguous, however, many database administrators and programmers use the term database to cover both meanings.

Many professionals consider a collection of data to constitute a database only if it has certain properties: for example, if the data is managed to ensure its integrity and quality, if it allows shared access by a community of users, if it has a schema, or if it supports a query language. However, there is no definition of these properties that is universally agreed upon.

Database management systems are usually categorized according to the data model that they support: relational, object-relational, network, and so on. The data model will tend to determine the query languages that are available to access the database. A great deal of the internal engineering of a DBMS, however, is independent of the data model, and is concerned with managing factors such as performance, concurrency, integrity, and recovery from hardware failures. In these areas there are large differences between products.

Database Models

Various techniques are used to model data structure.

Most database systems are built around one particular data model, although it is increasingly common for products to offer support for more than one model. For any one logical model various physical implementations may be possible, and most products will offer the user some level of control in tuning the physical implementation, since the choices that are made have a significant effect on performance. An example is the relational model: all serious implementations of the relational model allow the creation of indexes which provide fast access to rows in a table if the values of certain columns are known.

Flat Model

This may not strictly qualify as a data model, as defined above. The flat (or table) model consists of a single, two-dimensional array of data elements, where all members of a given column are assumed to be similar values, and all members of a row are assumed to be related to one another.

Hierarchical Model

In a hierarchical model, data is organized into a tree-like structure, implying a single upward link in each record to describe the nesting, and a sort field to keep the records in a particular order in each same-level list.

Relational Model

Three key terms are used extensively in relational database models: relations, attributes, and domains. A relation is a table with columns and rows. The named columns of the relation are called attributes, and the domain is the set of values the attributes are allowed to take.

The basic data structure of the relational model is the table, where information about a particular entity (say, an employee) is represented in columns and rows (also called tuples). Thus, the "relation" in "relational database" refers to the various tables in the database; a relation is a set of tuples. The columns enumerate the various attributes of the entity (the employee's name, address or phone number, for example), and a row is an actual instance of the entity (a specific employee) that is represented by the relation. As a result, each tuple of the employee table represents various attributes of a single employee.

All relations (and, thus, tables) in a relational database have to adhere to some basic rules to qualify as relations. First, the ordering of columns is immaterial in a table. Second, there can't be identical tuples or rows in a table. And third, each tuple will contain a single value for each of its attributes.

A relational database contains multiple tables, each similar to the one in the "flat" database model. One of the strengths of the relational model is that, in principle, any value occurring in two different records (belonging to the same table or to different tables), implies a relationship among those two records. Yet, in order to enforce explicit integrity constraints, relationships between records in tables can also be defined explicitly, by identifying or non-identifying parent-child relationships characterized by assigning cardinality (1:1, (0)1:M, M:M). Tables can also have a designated single attribute or a set of attributes that can act as a "key", which can be used to uniquely identify each tuple in the table.

A key that can be used to uniquely identify a row in a table is called a primary key. Keys are commonly used to join or combine data from two or more tables. For example, an Employee table may contain a column named Location which contains a value that matches the key of a Location table. Keys are also critical in the creation of indices, which facilitate fast retrieval of data from large tables. Any column can be a key, or multiple columns can be grouped together into a compound key. It is not necessary to define all the keys in advance; a column can be used as a key even if it was not originally intended to be one.

Relational Operations

Users (or programs) request data from a relational database by sending it a query that is written in a special language, usually a dialect of SQL. Although SQL was originally intended for end-users, it is much more common for SQL queries to be embedded into software that provides an easier user interface. Many web sites, such as Wikipedia, perform SQL queries when generating pages.

In response to a query, the database returns a result set, which is just a list of rows containing the answers. The simplest query is just to return all the rows from a table, but more often, the rows are filtered in some way to return just the answer wanted. Often, data from multiple tables are combined into one, by doing a join. There are a number of relational operations in addition to join.

Normal Forms

Relations are classified based upon the types of anomalies to which they're vulnerable. A database that's in the first normal form is vulnerable to all types of anomalies, while a database that's in the domain/key normal form has no modification anomalies. Normal forms are hierarchical in nature. That is, the lowest level is the first normal form, and the database cannot meet the requirements for higher level normal forms without first having met all the requirements of the lesser normal forms.

Object Database Models

In recent years, the object-oriented paradigm has been applied to database technology, creating a new programming model known as object databases. These databases attempt to bring the database world and the application programming world closer together, in particular by ensuring that the database uses the same type system as the application program. This aims to avoid the overhead (sometimes referred to as the impedance mismatch) of converting information between its representation in the database (for example as rows in tables) and its representation in the application program (typically as objects). At the same time, object databases attempt to introduce the key ideas of object programming, such as encapsulation and polymorphism, into the world of databases.

A variety of these ways have been tried for storing objects in a database. Some products have approached the problem from the application programming end, by making the objects manipulated by the program persistent. This also typically requires the addition of some kind of query language, since conventional programming languages do not have the ability to find objects based on their information content. Others have attacked the problem from the database end, by defining an object-oriented data model for the database, and defining a database programming language that allows full programming capabilities as well as traditional query facilities.

Post-Relational Database Models

Several products have been identified as post-relational because the data model incorporates relations but is not constrained by the Information Principle, requiring that all information is represented by data values in relations. Products using a post-relational data model typically employ a model that actually pre-dates the relational model. These might be identified as a directed graph with trees on the nodes.

Examples of models that could be classified as post-relational are PICK (a.k.a. “MultiValue”) and MUMPS.

Database Internals

Storage & Physical Database Design

Database tables/indexes are typically stored in memory or on hard disk in one of many forms, ordered/unordered Flat files, ISAM, Heaps, Hash buckets or B+ Trees. These have various advantages and disadvantages discussed in further in the main article on this topic. The most commonly used are B+trees and ISAM.

Other important design choices relate to the clustering of data by category (such as grouping data by month, or location), creating pre-computed views known as materialized views, partitioning data by range or hash. As well memory management and storage topology can be important design choices for database designers. Just as normalization is used to reduce storage requirements and improve the extensibility of the database, conversely denormalization is often used to reduce join complexity and reduce execution time for queries.

Indexing

All of these databases can take advantage of indexing to increase their speed, and this technology has advanced tremendously since its early uses in the 1960s and 1970s. The most common kind of index is a sorted list of the contents of some particular table column, with pointers to the row associated with the value. An index allows a set of table rows matching some criterion to be located quickly. Typically, indexes are stored in the same various techniques mentioned above (such as B-trees, hashes, and linked lists). Usually, a specific technique is chosen by the database designer to increase efficiency in the particular case of the type of index required.

Relational DBMSs have the advantage that indexes can be created or dropped without changing existing applications making use of it. The database chooses between many different strategies based on which one it estimates will run the fastest. In other words, indexes are transparent to the application or end user querying the database; while they affect performance, any SQL command will run with or without indexes existing in the database.

Relational DBMSs utilize many different algorithms to compute the result of an SQL statement. The RDBMS will produce a plan of how to execute the query, which is generated by analyzing the run times of the different algorithms and selecting the quickest. Some of the key algorithms that deal with joins are Nested loop join, Sort-Merge Join and Hash Join. Which of these is chosen depends on whether an index exists, what type it is, and its cardinality.

Transactions & Concurrency

In addition to their data model, most practical databases ("transactional databases") attempt to enforce a database transaction . Ideally, the database software should enforce the ACID rules, summarized here:

- Atomicity: Either all the tasks in a transaction must be done, or none of them. The transaction must be completed, or else it must be undone (rolled back).
- Consistency: Every transaction must preserve the integrity constraints — the declared consistency rules — of the database. It cannot place the data in a contradictory state.
- Isolation: Two simultaneous transactions cannot interfere with one another. Intermediate results within a transaction are not visible to other transactions.
- Durability: Completed transactions cannot be aborted later or their results discarded. They must persist through (for instance) restarts of the DBMS after crashes
- A cascading rollback occurs in database systems when a transaction (T1) causes a failure and a rollback must be performed. Other transactions dependent on T1's actions must also be rolled back due to T1's failure, thus causing a cascading effect.

In practice, many DBMS's allow most of these rules to be selectively relaxed for better performance.

Concurrency control is a method used to ensure that transactions are executed in a safe manner and follow the ACID rules. The DBMS must be able to ensure that only serializable, recoverable schedules are allowed, and that no actions of committed transactions are lost while undoing aborted transactions.

Replication

Replication of databases is closely related to transactions. If a database can log its individual actions, it is possible to create a duplicate of the data in real time. The duplicate can be used to improve performance or availability of the whole database system. Common replication concepts include:

- Master/Slave Replication: All write requests are performed on the master and then replicated to the slaves
- Quorum: The result of Read and Write requests are calculated by querying a "majority" of replicas.
- Multimaster: Two or more replicas sync each other via a transaction identifier.

Security

Database security is the system, processes, and procedures that protect a database from unintended activity.

Applications Of Databases

Databases are used in many applications, spanning virtually the entire range of computer software. Databases are the preferred method of storage for large multi-user applications, where coordination between many users is needed. Even individual users find them convenient, and many electronic mail programs and personal organizers are based on standard database technology. Software database drivers are available for most database platforms so that application software can use a common application programming interface (API) to retrieve the information stored in a database. Two commonly used database APIs are JDBC and ODBC.

Data & Networking Practises – Application Programming Interface (API)

What Is It?

An application programming interface (API) is a source code interface that a computer system or program library provides to support requests for services to be made of it by a computer program. An API differs from an application binary interface in that it is specified in terms of a programming language that can be compiled when an application is built, rather than an explicit low level description of how data is laid out in memory.

The software that provides the functionality described by an API is said to be an implementation of the API. The API itself is abstract, in that it specifies an interface and does not get involved with implementation details.

Two well known APIs are the Single UNIX Specification and the Microsoft Windows API.

An API is often a part of a software development kit (SDK).

The term API is used in two related senses:

- A coherent interface consisting of several classes or several sets of related functions or procedures.
- A single entry point such as a method, function or procedure.

Design Models

There are various design models for APIs. Interfaces intended for the fastest execution often consist of sets of functions, procedures, variables and data structures. However, other models exist as well - such as the interpreter used to evaluate expressions in ECMAScript/JavaScript or in the abstraction layer - which relieve the programmer from needing to know how the functions of the API relate to the lower levels of abstraction. This makes it possible to redesign or improve the functions within the API without breaking code that relies on it.

Some APIs, such as the ones standard to an operating system, are implemented as separate code libraries that are distributed with the operating system. Others require software publishers to integrate the API functionality directly into the application. This forms another distinction in the examples above. Microsoft Windows APIs come with the operating system for anyone to use. Software for embedded systems such as video game consoles generally falls into the application-integrated category. While an official PlayStation API document may be interesting to read, it is of little use without its corresponding implementation, in the form of a separate library or software development kit.

An API that does not require royalties for access and usage is called "open". Although usually authoritative "reference implementations" exist for an API (such as Microsoft Windows for the Win32 API), there is nothing that prevents the creation of additional implementations. For example, most of the Win32 API can be provided under a UNIX system using software called Wine.

In countries subject to software patents it is sometimes lawful to analyze API implementations to produce a compatible one. This technique is called reverse engineering for the purposes of interoperability. However, the legal situation is often ambiguous, so that care and legal counsel should be taken before the reverse engineering is carried out. For example, while APIs usually do not have an obvious legal status, they might include patents that may not be used until, and if, the patent holder gives permission. Many countries however do not grant software patents and within these countries only the protection of copyright applies to an API.

Release Policies

Two general lines of API publishing policies:

1. Some companies protect information on their APIs from the general public. For example, Sony used to make its official PlayStation 2 API available only to licensed PlayStation developers. This enabled Sony to control who wrote PlayStation 2 games. Such control can have quality control benefits and potential license revenue.
2. Some companies make their APIs freely available. For example, Microsoft makes most of its API information public, so that software will be written for the Windows platform.

It is to be expected that companies base their choice of publishing policy on maximizing benefit to themselves.

Some Example APIs

- The PC BIOS call interface
- Single UNIX Specification (SUS)
- Microsoft Win32 API
- Java Platform, Enterprise Edition APIs
- ASPI for SCSI device interfacing
- Carbon and Cocoa for the Macintosh OS
- OpenGL cross-platform API
- DirectX for Microsoft Windows
- Simple DirectMedia Layer (SDL)
- Google Maps API

Data & Networking Practises – Informix Dynamic Server

(Author's Note: The text in this section courtesy of Eric Clementi of IBM Atlanta)

Key Summary

Informix Dynamic Server (IDS) V9.4 is a best-of-breed, general-purpose OLTP database for the enterprise and workgroup.

- It is based on the parallel everything, fully multi-threaded Dynamic Scalable Architecture (DSA). This facilitates more efficient hardware resource utilization so you need less hardware to support your growing business needs.
- It uses object relational technology that dramatically increases the performance and efficiency of enterprise applications.
- Is a scalable database that dynamically adjusts core engine operational parameters, without degradation of performance, as requirements occur for larger databases or more concurrent users. This saves administrative time and resource as well as making more efficient use of hardware resources.
- Shows performance improvements of eight to fifteen percent over all previous IDS releases, based on internal performance testing.
- Removes many of the previous IDS size limitations. IDS V9.4 can now handle greater than 2 GB chunks of data, and much larger instance and file sizes than in previous releases. The net result is that up to 128 petabytes can now be stored in a single instance. (128 petabytes is equal to 12,800 Libraries of Congress!)
- Enhances application and data availability exponentially, using proprietary HDR and ER technologies to coexist within the same instance.
- Provides the highest levels of enterprise application availability while increasing transaction rates and response times. This enhances ease of management and decreases total cost of operation.
- Continues to support the latest industry standards including JDBC, ODBC, and OLE/DB.
- Offers industry standard security

Overview

Informix Dynamic Server (IDS) 9.4 is a best-of-breed online transaction processing (OLTP) database for enterprise and workgroup computing.

IDS 9.4 protects data assets in a highly-dependable database management system. High Availability Data Replication (HDR) provides complete turn-key disaster recovery. Enterprise Replication (ER) provides selective replication of data across multiple geographic locations.

IDS is built on Dynamic Scalable Architecture (DSA) that uses hardware resources more efficiently and minimizes hardware requirements. IDS 9.4 increases the maximum size of an IDS instance from 4 terabytes to a theoretical 128 petabytes -- enabling use of today's large disk drives.

IDS 9.4 simplifies and automates tasks traditionally associated with maintaining enterprise databases. Automated backup and restore functions eliminate many manual administration tasks.

IDS 9.4 provides increased flexibility and compatibility for Business Partner applications, including enhanced support of industry standard SQL syntax.

IDS 9.4 At A Glance

- Delivers the performance and reliability needed for always available, global, Web application-based OLTP
- Allows more business to be done with fewer resources, saving time and money
- Allows unlimited scalability through the removal of previous size limitations
- Maximizes data operations for the workgroup and the global enterprise through enhanced replication capabilities
- Provides the ease of use, self-management, and embedded features needed to support Business Partners and their next-generation applications
- Improves usability for developers and system administrators by including Server Studio JE from AGS to allow SQL Editing and Schema exploration

Key Features

IDS V9.4 includes the following features:

- Is based on the parallel everything, fully multi-threaded Dynamic Scalable Architecture (DSA) that guarantees you industry-leading performance while maximizing your resource utilization.
- Offers improvements in performance, availability, reliability, scalability, security, and SQL compatibility.
- Removes many of the previous size limitations making IDS an industry-leading scalable database that dynamically adjusts core engine operating parameters for optimal usage of system resources.
- The High Availability Data Replication and Enterprise Replication of IDS V9.4 can now coexist on key nodes within a worldwide network.

Product Versions

IDS V9.4 is available in three configuration options:

- **Workgroup Edition V9.4:** Ideal for small- to medium-size businesses and departmental solutions that require the power of an Enterprise Database.
- **Enterprise Edition V9.4:** Provides all of the features of IDS Workgroup Edition plus features that take advantage of today’s large-scale SMP systems. License is based on number of concurrent sessions.
- **Enterprise Edition Unlimited V9.4:** Same features as IDS Enterprise Edition except license is CPU-based.

Product Description

Informix Dynamic Server (IDS) V9.4 is a best-of-breed, general-purpose OLTP database for the enterprise and workgroup.

- Is based on the parallel everything, fully multi-threaded Dynamic Scalable Architecture (DSA). This facilitates more efficient hardware resource utilization so you need less hardware to support your growing business needs.

- Uses object relational technology that dramatically increases the performance and efficiency of enterprise applications.

- Is a scalable database that dynamically adjusts core engine operational parameters, without degradation of performance, as requirements occur for larger databases or more concurrent users. This saves administrative time and resource as well as making more efficient use of hardware resources.

- Shows performance improvements of eight to fifteen percent over all previous IDS releases, based on internal performance testing.

- Removes many of the previous IDS size limitations. IDS V9.4 can now handle greater than 2 GB chunks of data, and much larger instance and file sizes than in previous releases. The net result is that up to 128 petabytes can now be stored in a single instance. (128 petabytes is equal to 12,800 Libraries of Congress!)

- Enhances application and data availability exponentially, using proprietary HDR and ER technologies to coexist within the same instance.

- Provides the highest levels of enterprise application availability while increasing transaction rates and response times. This enhances ease of management and decreases total cost of operation.

- Continues to support the latest industry standards including JDBC, ODBC, and OLE/DB.

- Upgrading from IDS V7.x to V9.4 is simple and reliable.

More About Features For Various Versions

IDS V9.4 Workgroup Edition features and attributes include:

High Availability

- Point-in-time instance wide restore
- Table level backup and recovery
- XBSA support for easy integration of third-party backup tools
- Supports high-availability clustering solutions

High-Performance Technology

- Multithreaded architecture for SMPs with asynchronous I/O
- Cascading deletes; in-place table alters add and delete columns in seconds; bi-directional indices

Federated Data Access

- IBM distributed SQL supports queries across multiple IBM and third-party databases with IBM Enterprise Gateway products

IDS V9.4 Enterprise Edition has all the features and attributes of Workgroup Edition plus:

- **Data Partitioning** — Partition your data across multiple disks or disk subsystems via simple round-robin distribution or powerful range (expression)-based fragmentation techniques.

- **Parallel Data Query** — Once your data is partitioned, you can leverage IDS's parallel data query technology to query each partition in parallel. Use this feature to dramatically speed the queries in your applications.

- **Parallel Backup and Restore** — IDS Enterprise Edition can backup and restore your data in parallel. No need to serially backup your database. Instead, backup each partition individually. Online consistent backups.

- **Parallel Load and Unload** — The high performance loader, provided with IDS Enterprise Edition, takes advantage of IDS's data partitioning model to speed moving data in and out of your database.

- **Enterprise Replication** — A high speed, very flexible, low latency replication solution that is ideal for the needs of the enterprise.

- **High Availability Data Replication (HDR)** — HDR is IDS's proven high-availability solution. HDR has been shipping with IDS since 1993 and continues to provide the very high levels of continuous availability that Informix customers have enjoyed for nearly a decade.

Product Bundle

IBM Informix Dynamic Server V9.4 includes the following core products:

IBM Informix Connect V2.81: A runtime connectivity product that includes the runtime libraries of IBM Informix Client SDK. These libraries are required by applications running on client machines when accessing Informix servers. Informix Connect is needed when finished applications are ready to be deployed.

IBM Informix Server Administrator V1.5: A browser-based, cross-platform database server administration tool. It provides an easy-to-use interface for the IDS command line, thus eliminating the need to memorize commands and options.

IBM Informix Java Database Connectivity (JDBC) V2.21: The Java Soft specification of a standard API that allows Java programs to access database management systems. The JDBC API consists of a set of interfaces and classes written in the Java programming language. This JDBC driver also includes Embedded SQL/J which supports embedded SQL in Java. JDBC V2.21 is a platform-independent, industry-standard Type 4 JDBC driver that provides enhanced support for distributed transactions and is optimized to work with IBM Web Sphere Application Server. JDBC V2.21 promotes accessibility to IBM Informix database servers from Java client applications, provides openness through XML support, fosters scalability through its connection pool management feature, and supports extensibility with a user-defined data type (UDT) routine manager that simplifies the creation and use of UDTs.

Informix Dynamic Server 9.4 – More Information

- The best-performing IBM Informix® Dynamic Server (IDS) release available, exceeds all previous online transaction processing (OLTP) performance records

- Exceptional scalability to meet the unpredictable challenges of on demand e-business
- Upgrading to IDS 9.40 from any current IDS product is a simple, highly automated task
- Smooth, automated upgrade
- Increased flexibility and compatibility for business partner applications

Performance And Availability For Today's Continual Computing

IDS is a best-of-breed OLTP database for enterprise and workgroup computing. It has always supported very large database systems in demanding operational environments. In Version 9.40, this practice persists with exceptional performance, availability, scalability, security, simplified manageability and support of Business Partner applications. IBM continues longstanding traditions within IBM and Informix of delivering the mission-critical data infrastructure that enterprises need to run their businesses.

IDS 9.40 provides the essential qualities needed for today's always-available, Web-application based OLTP:

- **High performance**—With Version 9.40, IDS exceeds its already high OLTP performance by 8 to 15 percent with a more-efficient use of large amounts of memory and even less contention between concurrent users.
- **Data availability**—IDS High Availability Data Replication (HDR) provides complete turn-key disaster recovery. IDS Enterprise Replication (ER) provides selective replication of data across multiple geographic locations. IDS 9.40 enables both methods to coexist on key nodes within a worldwide network.

Scalability For Ever-Increasing Data

With databases getting larger and more concurrent users accessing them, IDS provides much-needed, exceptional scalability. It dynamically adjusts core engine operating parameters without performance degradation. At the same time, its dynamic, scalable architecture uses hardware resources more efficiently to minimize hardware requirements.

The enhanced scalability of IDS 9.40 increases the maximum size of an IDS instance from 4 terabytes to a theoretical 128 petabytes. In addition to improving performance, this scalability enables usage of today's large disk drives and eases administration overhead by significantly reducing the number of chunks that a database administrator (DBA) must manage.

Industry-Standard Security

IDS has always supported a set of open, industry-standard security mechanisms such as roles, UNIX® password-based authentication and database management system (DBMS) schema authorizations. IDS 9.40 also supports OpenSSL encryption libraries for communication security and Pluggable Authentication Modules (PAMs) for user authentication. These open standards ensure flexibility and maximum security with easier validation and verification. This IDS release also makes Lightweight Directory Access Protocol (LDAP) support much easier.

Increasing Productivity

IDS has a tradition of providing tools to help a small number of administrators manage a large number of systems. IDS 9.40 continues to reduce the complexity, time, DBA skills and cost associated with database management by eliminating, simplifying and automating many tasks typically associated with an enterprise-class database.

Upgrading to IDS 9.40 from any current IDS product is a simple, highly automated task. Also, the browser-based interface for IDS, Informix System Administrator (ISA), allows all common DBA activities to be administered from anywhere on the network. This interface now supports Informix Storage Manager (ISM), enabling DBAs to edit and configure events and alarms to respond quickly to unexpected events. Automated backup and restore functions offers DBAs high level data availability and restoration while eliminating many manual tasks. And new tape utilities help make maximum use of backup media by indicating how much data can be written to a compressing tape drive.

Flexibility For Application Development

IBM and IDS 9.40 help Business Partners improve their profitability, reduce their risk and deliver best-of-breed solutions to customers. IDS 9.40 has enhanced support of industry-standard SQL and continued support of the latest standards such as Java TM Database Connectivity (JDBC), Open Database Connectivity (ODBC) and OLE DB. In response, leading packaged-application vendors have committed to port to IDS 9.40 within a few months of its release.

IDS 9.40 also continues to foster application development in a worldwide marketplace. New features support partners in publishing applications in any language and creating multinational applications. It allows changing languages dynamically to truly support multilingual applications.

Integration With Leading Technology

IDS and IBM are a powerful combination. IBM offers complementary data management and software products that integrate with and support IDS including DB2® Information Integrator, Web Sphere®, Tivoli® and a variety of IBM tools. IDS 9.40 reflects IBM’s ongoing commitment to open standards and leading-edge technology support including Web services, Linux, autonomic computing and on demand e-business.

Features & Key Benefits

High-Performance Technology

- Cleaner indexes and improved memory management for more-efficient use of memory and processors.
- Enhanced buffer priority management for more-efficient use of large amounts of memory for today’s 64-bit operating systems.
- Smarter query optimization including user-defined type costing with new spatial data types
- Shortened instruction sets for common tasks
- Reduced contention between concurrent users

Scalability Previous Limitations Are Eliminated In IDS 9.40

- Data capacity of a single instance increased from 4 TB to 128 PB
- Maximum chunk size increased from 2 GB to 4 TB

- Maximum number of chunks increased from 2048 to 32 K
- Maximum LVARCHAR size increased from 2 KB to 32 KB
- DBMS utility file size limit increased from 2 GB to 17 billion GB

Security

- Secure over-the-wire encryption using the industry-standard OpenSSL encryption libraries
- Configurable user authentication mechanisms using Pluggable Authentication Modules (PAMs)

System management

- Informix Storage Manager support in the browser-based ISA
- Automated backup and restore functions
- Increased media efficiency by tape-handling utilities
- More flexible restore options

Application development

- SQL language enhancements
- Expanded Unicode support
- Multiple collation support
- Silent installation
- Continued support of IBM Informix 4GL, SQL, ODBC, JDBC, OLE/DB, SQLJ

Available on

IBM AIX®, HP-UX, Sun Solaris; Linux on Intel® processor-compatible servers, Linux for IBM zSeries servers; Reliant UNIX; SGI IRIX; Compaq Tru64; Microsoft® Windows® 2000, XP, 2003 (.Net)

Database Tools

There are a lot of tools available to Informix customers. This appendix is intended to provide an executive summary of what's available sorted into the following classes;

- Application Development tools
- Business Intelligence tools
- Database Administration tools
- Information Integration tools
- Performance tools
- Web Development tools

Informix Application Development Tools

- Four J's Business Development Suite – An advanced business oriented compiler for the rapid development of applications.

- Informix 4GL Family (Compiler, RDS, ID) – A comprehensive fourth-generation application development and production environment.

- Informix Client Software Development Kit (CSDK) – Single packaging of several application programming interfaces (APIs) needed to develop applications for Informix servers.
- Informix Connect – A runtime connectivity product.
- Informix ESQL/C – The convenience of entering SQL statements directly into the C language source code.
- Informix ESQL/COBOL – An SQL application programming interface (SQL API) that lets you embed SQL statements directly into COBOL code.
- Informix JDBC – The Java Soft™ specification of a standard API that allows Java™ programs to access database management systems (DBMSs).
- Informix SQL – A database application development system that provides speed, power and security.

Business Intelligence Tools

- Informix MetaCube – Helps you drive smarter decisions out of your data quickly and incisively.
- IBM Office Connect - Provides front-end tool for databases as add-in to Microsoft Excel

Database Administration Tools

- Informix Server Administrator – A browser-based, cross-platform database server administration tool.
- Server Studio JE – An enterprise-level, Integrated Development and Management Environment (IDME) tools deployment platform developed by AGS.

Information Integration Tools

- DB2 Information Integrator – Information Integrator provides integrated, real-time access to diverse data as if it were a single database, regardless of where it resides.
- Informix Enterprise Gateway Manager – A SQL-based gateway that allows a customer's Informix tools, applications, and databases to interoperate transparently with non-Informix databases.
- Informix Enterprise Gateway Manager with DRDA – A UNIX®-based connectivity solution for IBM relational databases.

Performance Tools

- Informix MaxConnect – Improves system scalability and performance by increasing the number of users that can simultaneously connect to an Informix database server.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

- Informix I-Spy – A smart data warehouse monitoring and optimization tool designed for IBM Informix databases.

Web Development Tools

- Informix Data Director for Web – Provides a model-driven development environment designed explicitly for creating powerful database applications that can grow with your business.

- Informix Web DataBlade Module – A collection of tools, functions, and examples that ease development of "intelligent", interactive, Web-enabled database applications.

Data & Networking Practises – DataBlades

Introduction

The technology of DataBlades is extremely powerful but in many cases it's complex and difficult to get to grip with if you're not an Database Administrator or Business Analyst. We will begin by providing a description of the technology by means of an analogy.

The Experts Analogy

A way to understand how DataBlades work is to view them like people, specifically experts brought in to provide an organisation with specific expertise.

Let's say you were building a factory. You already have a company with staff that can do the logistical planning and finance, but you don't have an architect, engineer or builder on the staff to build the factory with.

So to build the factory you acquire the expertise by means of experts, in this case an architect, engineer and builder, to help you build the building.

DataBlades are very much like these experts brought on board. They live on the database directly and bring specific expertise, such as image or video management, to the enterprise. Luckily, unlike real experts they do not require a salary, once their on-board they're always there!

While the Project to build the factory is moving forward it hits a snag and requires another expert, in this case an electrician to wire the building. In the real world you simply hire the skills you need and the same applies to DataBlades, you simply deploy the new functionality you require on the database and carry on.

DataBlades are pretty much like organisations that employ many experts, they end up with huge reservoirs of skills at their disposal giving the organization robustness, flexibility and far more options than a skills poor company.

The core idea behind DataBlades is to make expertise available directly onto the database, as opposed to all manner of strange patchwork architectures, making it easy to manage and simple to exploit.

Making A Better Team

In any organization employing experts makes sense if you have specific needs, but how these experts work with the rest of the company and with each other is quite important. The same guidelines that apply to people apply well to DataBlades.

Though there are several similar technologies available in the marketplace, it is very important that they

- co-exist
- coordinate
- integrate
- synchronize

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited*****.io***

- harmonize

To be more specific this means they must, just like people in a team;

- **Co-exist** – meaning there can be no exclusive combinations, i.e. this one does this particular thing well but cannot work with this other one that does the same or a different thing well and so one
- **Integrate** – this means that the experts have to be a part of the company to be effectively managed and in the case of DataBlades mean they have to be directly deployed on top of the database (cost and complexity saving)
- **Coordinate** – they need to be able to exchange tasks and communicate with each other
- **Synchronize** – they need to be able to co-operate on many levels and need to be able to be managed centrally from one point
- **Harmonize** – they need to speak the same language and the sum of the whole must be greater than the parts

IBM Informix DataBlades

The solution design not only calls for a world class OLTP and Object-relational database, it also requires some far more sophisticated functions. The DataBlades provide that functionality.

DataBlades are set of function sets which run directly integrated in the database. Where traditional programs have an intelligence layer between the client and the server or even reference other servers, DataBlades allow you to deploy complex functions directly in the database eliminating much code customisation and providing high transaction throughputs.

Where Do They Come From?

There are dozens of DataBlades. The major DataBlades are produced by IBM, but many 3rd party vendors (such as Adobe and ESRI) and IBM Informix customers (such as the U.S. Navy and N.A.S.A.) have written their own specifically for the Informix platform.

Though relatively easy to write, meaning the method to create new DataBlades from scratch is relatively simple using either Java or C, many specialist blades with highly sophisticated business intelligence and exotic functions have been written.

One example of a specialized DataBlade is the NAG TimeSeries DataBlade. The National Algorithms Group (NAG) produced a highly sophisticated DataBlade in which hundreds of stock market specific algorithms are stored. If the Informix customer now requires to develop a market financial application, his or her programmer's can call this business logic straight from the database as opposed to coding it themselves. The time saving on development is tremendous.

Combinations

The DataBlades can be combined to offer hundreds of standard versions of the product in days, enabling nearly any organization rapidly and eliminating costly and timeous customisations.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

What Is DataBlade Technology?

- Extends server functionality
- Fuels high-speed content management
- Empowers customization
- Manages multimedia content with finesse
- Captures geospatial business intelligence.

Leveraging Data Across The Internet

The opportunities afforded by the Internet are equalled only by the data-driven challenges it presents:

How to extend the reach of your enterprise to the Internet and still manage costs? How to transform bits and bytes of business data into usable content? How to leverage data as business intelligence, making your company more competitive?

Internet applications are about content as well as data. Today, content is derived from a wide variety of sources – from audio and video to text and images – and its potential is as exhaustive. However, the cost and complexity of managing and integrating data and content for your Web-based applications can be a daunting challenge for your enterprise.

IBM Informix: Rich Content Management

IBM Informix® software can help solve these mission-critical concerns. Start with an inherently extensible application foundation – IBM Informix Internet Foundation – and add the rich content management capabilities of IBM Informix DataBlade® modules. IBM Informix DataBlade technology integrates any data from any source to provide a proven, stable, and flexible Internet application environment. IBM Informix DataBlade modules put the power of the Internet directly into the IBM Informix Dynamic Server. It's that simple.

Superior Server Extensions

IBM Informix DataBlade modules actually extend the power of IBM Informix Dynamic Server to suit the unique requirements of your business. DataBlade modules are not just options but actual server extensions integrated into the very core of the engine, delivering unbelievable application functionality and superior performance.

Imagine combing a database for all residences which are located within 200 meters of a fire hydrant. Using standard SQL, this seemingly simple query becomes a hugely complicated and time-consumptive task. With IBM Informix DataBlade technology, you can enrich the intelligence of the application by adding geospatial and regional information – as a natural extension to the data managed by the server. The result? Fewer lines of code, faster execution and a more exact answer rather than an approximation.

DataBlade modules have been enterprise-tested in complex mission-critical applications for years. Consider these cross-industry examples:

- Major telecommunications providers have leveraged DataBlade technology to add spatial intelligence and visual mapping to production systems – for use in network management and planning.

- Large financial institutions use the IBM Informix Time Series DataBlade module to track daily trading information and help with analyzing and predicting market trends.
- The wireless telecommunications industry has combined the IBM Informix Spatial DataBlade and Time Series DataBlade modules for fraud-detection purposes.
- Content providers and publishers are using IBM Informix Web DataBlade technology for dynamic Web-page creation, automatic content importation, and highly flexible content organization and output options – with unsurpassed scalability and reliability standards –many enabling more than 6 million page views per month.

DataBlade modules may be used separately or integrated as a suite to create a unique information management solution to suit your organizational needs. Across industries, in the face of change, encompassing all kinds of content – IBM Informix DataBlade technology is the smarter business choice for efficient, effective and distinctive content management.

High-Speed Content Management, End-To-End

IBM Informix DataBlade modules give you the tools you need to intelligently manage rich, diverse data. DataBlade modules integrate traditional alphanumeric data types with rich content – without sacrificing the reliability and scalability of the traditional relational DBMS.

IBM Informix DataBlade technology lets businesses treat Web sites as applications, and enables the IBM Informix Internet Foundation to manage and dynamically deliver all site content. The server can be rapidly modified to accommodate new data types as business requirements evolve. With IBM Informix DataBlade technology, you manage complex situations quickly and cost effectively right out of the gate, giving you a consistently faster time to market – and a big jump on your competition.

DataBlade Technology: Working With You And For You

DataBlade technology empowers your customization efforts; you can embed your technology and expertise as an integral part of the IBM Informix Internet Foundation. While pre-built DataBlade modules for text and Web document management are bundled with IBM Informix Internet Foundation, we also include a full development kit for building your own specialized DataBlade modules. IBM provides a single API development kit for Java TM, C, C++, J++, or stored procedure language (SPL). You also are given the opportunity to choose from a portfolio of third-party DataBlade modules.

Managing Content Across Mediums

To publish compelling Internet content you need a database that can easily manage large volumes of digital information, including images, audio and video. Digital media DataBlade modules manage, publish, convert and efficiently store all of your company's digital data with finesse.

Likewise, developing usable content requires navigation through, and conversion of, large amounts of unstructured text. With IBM Informix Text DataBlade modules, you can search large amounts of text quickly, thereby extracting the true value inside hidden information resources. DataBlade modules can publish manuals, reports, formatted documents, e-mail, Web pages, faxes, PDF files and presentations to the Internet – quickly, with promised quality at a low cost. Plus, with DataBlade technology, documents can be stored directly in the database, creating a universal repository, which eliminates the need for separate file systems and allows full concurrency, recovery and transaction control.

Leveraging Geospatial Knowledge

Organizations across industries are realizing the advantages of leveraging the wealth of their geographic information.

Why are certain products selling better in particular markets? How can customer preferences be identified by region? Where are the untapped market niches?

In the global marketplace, it is increasingly important to capture meaningful answers to questions like these.

IBM Informix, along with industry-leading DataBlade modules from third parties such as MapInfo and Environmental Systems Research Institution, Inc. (ESRI), allows organizations to intelligently manage complex geospatial information alongside traditional data – without sacrificing the efficiency of the relational database model. Just another way IBM Informix DataBlade technology helps you access and exploit your organization's information assets – anytime.

Varieties

There are dozens of DataBlades available from various accredited vendors to customers running Informix databases. Some of these have some exotic and industry specific functions, i.e. the Russian DataBlade which supports true grammatical translations to and from Russian.

In this section we will focus in the major, or most prominent, DataBlades available and offer a quick summary of their key features.

The Major DataBlades

- C-ISAM Blade
- Excalibur Image Blade
- Excalibur Text Blade
- Geodetic Blade
- Image Foundation
- Spatial Blade
- TimeSeries Blade
- TimeSeries NAG Blade
- TimeSeries Real-Time Loader
- Video Foundation
- Web Blade

C-ISAM Blade

Overview

- IBM Informix® C-ISAM® DataBlade® module is a library of C functions that efficiently manages Indexed Sequential Access Method (ISAM) files.
- Developed specifically to help customers and software vendors who are interested in adding RDBMS features to their C-ISAM environment or are working toward migrating C-ISAM applications to an RDBMS environment.
- Industry-standard ISAM product for UNIX® in the United States and Europe for over eight years

- Fast data access
- Reliability
- Flexible indexing options
- Large file support
- Compatibility
- A SQL Access component which provides an SQL interface to the C-ISAM data.
- A Server Storage component which provides the capability to store ISAM data directly in the database server while allowing C-ISAM programs to continue accessing this data.

Excalibur Image Blade

Overview

- The Excalibur Image DataBlade module combines Excalibur image technology with the Informix® Dynamic Server with Universal Data to store, retrieve, and search images in a database.
- Based on technology from Excalibur Technologies Corporation, and is co-developed and co-supported by Informix and Excalibur.
- Multiple image file formats are supported to address needs of specific industries.
- Excalibur's Adaptive Pattern Recognition Processing (APRP) enables image-based searching of visual data by colour, shape or texture.
- Utilizes the indexing capabilities Excalibur's Visual Retrieval Ware with the Informix Dynamic Server to provide optimal query performance
- The image technology contained in the Excalibur Image DataBlade module is based on feature-vector extraction search techniques. It allows you to perform content-based searches of groups of images by reference to the binary patterns in feature vectors.

Excalibur Text Blade

Overview

- Informix® Excalibur Text enables full text search capabilities so you can take full advantage of an engine that is optimized for indexing and searching text information, using proximity searches and other features. Customers can now add extensive text-searching to many existing applications.
- Full Text Indexing, including extensive support for fuzzy-search logic, especially important when indexing scanned text
- Rapid query results
- Current support for documents in ASCII, Microsoft Word®, Excel®, PowerPoint®, HTML, PDF, Word Perfect, and other formats
- Adaptive Pattern Recognition Process engine (APRP) for maximum precision, recall, and relevancy in applications that require either full text retrieval or ad hoc text query
- Feature-rich (multiple stop-word lists, proximity searching, synonym lists

Geodetic Blade

Overview

- Combined with the IBM Informix Dynamic Server, the IBM Informix Geodetic DataBlade module provides you with the ability to manage geospatial information

referenced by latitude-longitude coordinates -- supporting global space- and time-based queries without limitations inherent in map projections.
- Geodetic DataBlade module treats the earth as a globe, not a flat map.
- No edges and no distortions
- Includes a time dimension and an altitude range in its spatial types
- Yields tremendous performance benefits where time and other criteria are often used in queries alongside the spatial criteria.
- A single index scan can cover all of these dimensions simultaneously.

Image Foundation

Overview

- The Informix® Image Foundation DataBlade® provides a base on which new or specialized image types and image processing technologies can be quickly added or changed.
- Open, secure , and scalable, providing a clear path toward reusing and repurposing valuable image assets.
- Store and retrieve images in the database or on remote computers and storage servers
- Store and retrieve image metadata
- Transform images using the industry-standard CVT command set
- Increases productivity and enhances workflow by allowing images to be automatically manipulated and deployed to applications

Spatial Blade

Overview

- The IBM Informix® Spatial DataBlade® brings all the significant features and benefits of the IBM Informix Dynamic Server™ (IDS) to location-based data. Many key business decisions involve location and proximity. The Spatial Datablade enables organizations to transform both traditional and location-based data into important information to help gain a competitive advantage.
- Informix Spatial DataBlade Module is an extension to the IBM IDS object-relational database engine which provides SQL-based spatial data types and functions, which can be used directly through standard SQL queries or with client-side GIS software such as that from ESRI and MapInfo.
- Delivers innovative spatial technology via a convenient no-charge download
- Generates vital business intelligence for a competitive edge
- Maximizes spatial data capabilities; enables critical business decisions
- Supports Enterprise Replication of spatial data types
- Drives industry-leading spatial query performance with the IBM IDS
- Extends database power
- Enables organizations to intelligently manage complex geospatial information alongside traditional data without sacrificing the efficiency of the relational database model.
- Built-in Informix R-tree multi-dimensional index to provide industry-leading spatial query performance
- Spatial family also includes the IBM Informix Geodetic DataBlade Module providing customers with the ability to manage geospatial information referenced by latitude-longitude coordinates.

TimeSeries Blade

Overview

- The IBM Informix TimeSeries DataBlade® module greatly expands the functionality of the database by adding sophisticated support for the management of time-series and temporal data.
- A time series is any set of data that is accessed in sequence by time which can be processed and analyzed in a chronological order.
- TimeSeries is a native type and the supplied functions are first-class functions that can appear anywhere in a SQL statement.
- Simplifies design and coding; and enables ad-hoc complex queries on time series, other complex data (e.g., text or spatial) and regular relational data, using SQL-3.
- Data is stored in a single, DSA-based DBMS.
- The IBM Informix NAG DataBlade module uses powerful algorithms from the Numerical Algorithms Group to perform complex calculations on TimeSeries data in the database engine.

TimeSeries NAG Blade

Overview

- Informix® Dynamic Server, TimeSeries and the NAG Financial DataBlade Module is uniquely positioned to provide a tick database.
- Uses powerful analytical algorithms from The Numerical Algorithms Group to operate directly on the data in the database engine.
- The Informix-NAG Financial DataBlade module moves the analysis closer to the data.
- Partnered with the IBM Informix TimeSeries DataBlade Module, the Informix NAG DataBlade Module can track daily trading information for forecasting, analyzing, and predicting market trends.

TimeSeries Real-Time Loader

Overview

- IBM Informix® TimeSeries Real-Time Loader ® is a data loader that works in conjunction with IBM Informix NAG DataBlade technology to achieve greater analytical performance than is possible with either traditional relational databases or standalone real-time analysis software.
- Specifically designed to load time-stamped data and make the data available to queries in real-time
- Runs on the IBM Informix Dynamic Server and leverages the IBM Informix NAG and IBM Informix TimeSeries DataBlade modules
- Provides instant visibility to real-time data
- Analyzes complex data and makes well-informed decisions gives organizations a competitive advantage

Video Foundation

Overview

- The Informix® Video Foundation DataBlade module is an open and scaleable software architecture that allows strategic third-party development partners to incorporate specific video technologies such as video servers, external control devices, compression codecs, or cataloguing tools into complete database management applications with the INFORMIX-Dynamic Server.
- Extends the capabilities of Informix database server to manage video content and metadata, or information about the content
- Metadata elements are stored in the database while the actual video content can be maintained on disk, video tape, video server, or other external storage devices.
- Storage capability addresses the "analog reality" of many video production facilities.

Web Blade

Overview

- The Informix® Web DataBlade module is a collection of tools, functions, and examples that ease development of "intelligent", interactive, Web-enabled database applications.
- The Web DataBlade module supports most Web Server Application Programming Interfaces (APIs), and enables a truly interactive Web site.
- Webdriver database client application build SQL queries that execute the WebExplode function for retrieving AppPages from your database.
- Enables you to customize Web applications using information from its configuration file, or stored in the database-- without gateway programming.
- Allows you to track persistent session variables between AppPages.
- NSAPI Webdriver component: This implementation of Webdriver is written with the Netscape Server API and is used only with Netscape Web servers.
- Apache Webdriver component: This implementation of Webdriver is written with the Apache API and is used only with Apache Web servers.
- ISAPI Webdriver component: This implementation of Webdriver is written with the Microsoft® Internet Information Server API and is used only with Microsoft Internet Information Web servers.
- CGI Webdriver component: This implementation of Webdriver is a standard CGI program that can be executed by all Web servers.
- The WebExplode Function is an SQL function that builds dynamic HTML pages based on data stored in your database.

Other DataBlades

There are many DataBlades available to the Informix database, so many that in fact it is not practical to list them all in this document. Not only are the abilities they can give an Informix database mind-boggling, because of their vast diversity the practical applications are endless.

We have instead opted to provide summaries on the most prominent other DataBlades after the major ones reviewed earlier in this document.

There is some common jargon used in the Executive Summaries below, they are:

- **UDR** = User Defined Routine
- **UDT** = User Defined Type

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

DataBlade	Description
BWA UDT DataBlade	The Business Web Application (BWA) user-defined type (UDT) identifies products on an e-commerce site. BWA is an intelligent part number, which means that it is composed of multiple parts, each of which has meaning. The BWA UDT is implemented as an opaque type and packaged as a DataBlade module. The DataBlade module includes C source code for implementing the opaque type support functions, btree index support, and user-defined routines (UDRs) that access each part separately.

DataBlade	Description
Chinese DataBlade	The Chinese DataBlade is a collection of data types and functions that extends INFORMIX Universal Server to enable you to input, index and search Chinese documents. This DataBlade provides the following data types for handling double-byte Chinese characters: CHIchar, CHIvarchar & CHItext. CHIchar is the equivalent of char in native IUS. It is 2 bytes long and represents one Chinese character. CHIvarchar is the equivalent of varchar in native IUS. Its maximum length is 255 bytes, accommodating a maximum of 127 Chinese characters. CHItext is the equivalent of text in native IUS. Its maximum length is 2**31 bytes, accommodating a maximum of 2**30 Chinese characters.
Exec DataBlade (exec_sql_udr)	The Exec DataBlade is a UDR that provides dynamic SQL functionality within an SPL procedure. It is a pair of user-defined functions -- a simple UDF and an Iterator -- that take arbitrary SQL statements, execute them, and return a result to whatever invoked it in the first place. Often, it is desirable to execute a SQL query that is generated at run-time within the ORDBMS. For example, a developer may not know the name of the temporary table they wish to run the query against, or they might want to append predicates to a query. In external programs, this can be accomplished using the ESQL/C SQLCA and DESCRIPTOR facilities. Unfortunately the INFORMIX Stored Procedure Language (SPL) does not support dynamic SQL. Queries must be hard-coded into the SPL logic. The objective of the Exec DataBlade is to remedy this. Exec consists of some user-defined functions (UDFs) that take a SQL query as an argument, execute it, and return a result (the format of which varies depending on the function and the kind of query). The Exec functions can handle most Data Definition Language (DDL) statements, and all Data Manipulation Language (DML) queries.
Grid DataBlade	The DataBlade is designed to handle 4-Dimensional (and 5-Dimensional) grids as a standard. Grids of lesser dimensionality are handled by taking one or more of the grid dimensions to be 1. It stores grids using a tiling scheme in conjunction with SmartBLOBs, with user control over the tile size. This allows very efficient generation of data products that involve only a small portion of the data in the overall grid (e.g., extracting planar oblique slices from volumetric grids, or retrieving 1D probes or sticks from 3D or 4D grids). It can also store the data in, and convert it between, more than 40 different planar mapping projections and often provides more than 50-fold increases in speed of data product generation compared to the conventional approach that does not involve tiling or SmartBLOBs. It is typically used in oceanographic and geophysical research.

DataBlade	Description
GridExtras DataBlade	The GridExtras DataBlade is a DataBlade developed by the U.S. Navy that extends the Grid DataBlade with additional complex routine functions, enabling developers to make direct calls to the functions instead of attempting to develop or deduce them outside of the database. Grid data is necessarily massive and complex, this greatly optimises the application architecture and performance.
HFD-EOS DataBlade	Hierarchical Data Format (HDF) is the standard data format for all NASA Earth Observing System (EOS) data products. It was developed to assist users in the transfer and manipulation of scientific data across diverse operating systems and computer platforms, using FORTRAN and C calling interfaces and utilities. HDF-EOS supports three geospatial data types (grid, point, and swath), providing uniform access to diverse data types in a geospatial context.
iUtil DataBlade	Differences in database vendor SQL support make it difficult to migrate applications between different database vendor products, especially if an application relies on a SQL feature in one database that is not available in the other. The IUtil component was developed to help a customer migrate an application from IBM DB2 to Informix IDS. It implements a specific subset of the proprietary functions that are built into DB2, and it can be extended to include virtually any others that you may require.
JPEG Image DataBlade	JPEG (JFIF) is one of the most popular image formats used on the web. This JPEG image DataBlade lets you store jpeg images together with extracted image properties in a database. Its user-defined image row type (JPGImage) and image processing routines make it suitable for storing, manipulating and searching jpeg images for the web as well as other image database applications. The JPGImage DataBlade is a plug-in module on Informix Dynamic Server. It extends the server's functionality through user defined image data types and routines to allow you to store, manipulate and retrieve JPEG images in a database more efficiently.
MIO DataBlade	The Multirepresentational lvarchar Opaque (MIO) DataBlade creates the idn_mrLvarchar opaque type, which stores character data up to 2 gigabytes. It is a multirepresentational (multirep) type, which means that small values get stored in-row and large values get stored in a smart blob.
MPEG DataBlade	The MPEG-2 DataBlade is used for structured video stream and meta data management and offers the integration of the MPEG video format into the database model and

DataBlade	Description
	schema of the Informix Universal Server (IUS).
Multi-Dimensional Spatial DataBlade (mdspatial)	This Spatial DataBlade module, based on both the Informix Shapes2 DataBlade module and on Rtree software developed by the original author, provides multidimensional spatial (opaque) types for the Informix Universal Server. It also provides the methods required by Informix to create and maintain spatial (rtree) indexes, and various cast and type constructor functions.
Node DataBlade	The Node data type is intended to resolve one of the hardest problems in relational databases; transitive closure. It doesn't resolve the problem completely, but it does get us 80% of the way. The transitive closure problem is endemic to data management problems, and not particularly well addressed by the relational model. The same basic problem is found modelling organizational hierarchies, networks, manufacturing and process control databases. There are a couple of optimisations for addressing the transitive closure problem, dealt with at length in Joe Celko's "SQL For Smarties" (Morgan Kaufmann, 1995). However, all of these optimisations introduce one of several kinds of anomalies, requiring that the database change more than one tuple in response to more than one fact. Also, the SQL needed to write the operations over the tree are complex, difficult to optimise and execute and -- even at their very best -- slow. The Node data type provides a quite lateral solution to this problem. The basic idea is to create a type and operations over that type which models the tree structure itself and to represent the facts of the tree structure in that way, rather than reduce these facts to overly simple relations.
Period DataBlade (period)	The Period DataBlade provides SQL support for managing information about fixed intervals in a timeline: real world phenomenon like the duration of a hotel stay, or the scheduled activity of some manufacturing equipment. The DataBlade implements a pair of OPAQUE TYPE User-Defined Type (UDT) instances, and a set of User-Defined Functions (UDFs) to perform the operations appropriate to these types. In addition to being a useful bundle of extensions in its own right, the Period DataBlade also provides a good example of several extensibility features: building an OPAQUE TYPE, how to use ORDBMS the Operator Class feature with R-Trees, a User-Defined Aggregate example, and a set of statistics and selectivity facilities.
Random DataBlade (random_udr)	This DataBlade implements several Random number generation UDRs. The idea is to be able to generate random numbers that follow several different distributions; Uniform, Binomial and Normal (or Gaussian.) These are

DataBlade	Description
	particularly useful generating large data sets for testing, and for certain simple kinds of sampling. The Blade consists of a number of UDRs (user-defined routines) that are designed to return sequences of random variables that comply with one or another distribution function. The most common such distribution is simply Uniform, which means the random function returns values in a range, all outcomes equally probable. Another set of UDRs deal with Binomial distributions (sequences of trials of fixed probability and Boolean result) and Normal (or Gaussian) distributions (distribution around a mean of known variance).
RGBA DataBlade	This DataBlade is intended for users that need to implement a sophisticated, fast and robust Colour Search and matching system. Examples of these include but are not limited to Image Search engines that want to retrieve images by comparing their contents. This DataBlade is also successfully used in Radar, Thermo, Microwave and Infra-red Image analysis. Colour handling solves important problems when dealing with image or video data. It handles the extracted data out of the information that is presented in the different parts of the otherwise overwhelming information. It enables others to improve the handling of base information that is required to uniquely store key information about higher level information. (An RGBA value is made up of four computer colour components: red, green, blue, and alpha. The red, green and blue components store the intensity for each colour as a value between 0 and 255. The alpha value is similarly stored as a value between 0 and 255 and represents the opacity of the colour represented.)
Russian DataBlade	The Russian Text DataBlade module gives the opportunity to build an index on a set of Russian language documents and perform high speed retrieval with boolean conditions on words, distance conditions, exact phrase conditions. Query may include options to perform search with substitutions and transpositions of letters in words. Russian language morphology available. RussianText DataBlade Module is a data types and functions library, which allow applications to perform more complex and high speed keyword searches then ordinary SQL.
Shapes DataBlade	The DataBlade creates several opaque types for managing spatial data. It implements a type hierarchy by using opaque types and implicit casts. It also uses R-tree support functions for variable-length opaque types.
Smart Blob Information DataBlade	The oncheck utility lets you monitor smart blob spaces (sbspaces) and monitor the smart blobs stored in those sbspaces, but it is difficult to associate oncheck output with a specific smart blob. The Smart Blob Information

DataBlade	Description
	DataBlade creates smart blob metadata and constructor UDRs to make it easier to monitor a specific smart blob and to create/update test smart blobs. The Smart Blob Information DataBlade metadata UDRs extract storage information about a smart blob.
Soundex DataBlade	Allows searches on phonetic sound or closest match to phonetic sound of a search string to data without introducing redundant columns into tables to hold the soundex code or compromising the relational integrity of the database. The Soundex DataBlade creates a new datatype which replaces char and the properties of phonetic identification become part of the behaviour of the datatype. This functionality is usually built into customer search screens, or street name search screens because spelling is often subjective, and mistakes in annotation are common.
Split DataBlade (split_udr)	This Blade contains two user-defined routines (UDRs) which split a single string into a COLLECTION of subvalues based on a delimiter. This is similar to the PERL operator of the same name, and has a similar purpose. The first of these routines is Split (LVARCHAR, LVARCHAR). It returns a single COLLECTION (LIST) of LVARCHAR instances which correspond to the sub-tokens in the first argument delimitered by any of the characters in the second. The second UDR is called ISPlit(LVARCHAR, LVARCHAR). It is an example of a 'C' iterator function. Like SPL functions that RETURN WITH RESUME, a 'C' iterator returns a series of values. The purpose of this function is to allow folk to iterate through the elements of the LIST more efficiently than they otherwise would.
SqlLib DataBlade	Differences in database vendor SQL support make it difficult to migrate applications between different database vendor products, especially if an application relies on a SQL feature in one database that is not available in the other. The SqlLib DataBlade module adds SQL routines that the Informix Dynamic Server (IDS) does not currently support natively but that are available in other vendor databases. Two source code implementations, one in C and the other in Java, are available to provide means to developers who wish to add the required functionality to IDS.
TerrainWorks DataBlade	This blade is used for the management of terrain-embedded linear networks. The TerrainWorks DataBlade has been designed to provide spatial data management professionals with a core tool for modelling, maintaining and querying large, seamless, linear networks that may be embedded on a terrain surface. Hydrological networks, road networks and electric power grids are all examples of this type of embedded network. To achieve truly seamless coverage, all geographic data 3D elements are stored as

DataBlade	Description
	(Latitude, Longitude, Height), with computation performed on the ellipsoid. Projection to a particular mapping plane is done ‘on the fly’, when required. The basic TerrainWorks DataBlade data objects, which include versioning and accuracy information at the feature level, are point, line and area features, partitioned coverages, linear networks and datachips (an abstract data type designed to store data in a very compact form that also supports fast access).
Verity Text Search DataBlade	The VTS DataBlade module enables Universal Data Option customers to use the Verity search engine to index and search text data stored in a variety of languages and formats in non-ANSI databases. As with standard database searches, customers use SQL statements for queries. The VTS DataBlade module performs both index and nonindex searches. The optimiser determines whether to use an index search by estimating cost efficiency. With the VTS DataBlade module, you can search for words and phrases, use expressions as search criteria, do proximity matching, and search for synonyms. Any search option provided by the Verity query language is available to you with the VTS DataBlade module.
VRML DataBlade	Virtual Reality Modelling Language (VRML) DataBlade is an extension module for an object-relational Informix Database Management System which allows the user to use data types and functions for the management of VRML scenes directly from the SQL layer. These functions include manipulation, composition and I/O of VRML scenes inside a database.

Other – Design Ergonomics

What Is It?

Ergonomics (or human factors) is the application of scientific information concerning humans to the design of objects, systems and environment for human use (definition adopted by the International Ergonomics Association in 2007). Ergonomics is commonly thought of as how companies design tasks and work areas to maximize the efficiency and quality of their employees' work. However, ergonomics comes into everything which involves people. Work systems, sports and leisure, health and safety should all embody ergonomics principles if well designed.(International Ergonomics Association in 2007) It is the applied science of equipment design intended to maximize productivity by reducing operator fatigue and discomfort. The field is also called biotechnology, human engineering, and human factors engineering.

Ergonomic research is primarily performed by ergonomists who study human capabilities in relationship to their work demands. Information derived from ergonomists contributes to the design and evaluation of tasks, jobs, products, environments and systems in order to make them compatible with the needs, abilities and limitations of people (IEA, 2000).

Ergonomic Domains

The IEA divides ergonomics broadly into three domains:

- Physical ergonomics deals with the human body's responses to physical and physiological stress.
- Cognitive ergonomics, also known as engineering psychology, concerns mental processes such as perception, attention, cognition, motor control, and memory storage and retrieval as they affect interactions among humans and other elements of a system. Relevant topics include mental workload, vigilance, decision making, skilled performance, human error, human-computer interaction, and training.
- Organizational ergonomics, or macroergonomics, is concerned with the optimization of sociotechnical systems, including their organizational structures, policies, and processes. Relevant topics include shift work, scheduling, job satisfaction, motivational theory, supervision, safety culture, teamwork, telework and ethics.

Foundations

Ergonomics draws on many disciplines in its study of humans and their environments, including anthropometry, biomechanics, mechanical engineering, industrial engineering, industrial design, kinesiology, physiology and psychology.

Typically, an ergonomist will have a BA or BS in Psychology, Industrial/Mechanical Engineering or Health Sciences, and usually a MA, MS or PhD in a related discipline. Many universities offer Master of Science degrees in Ergonomics, while some offer Master of Ergonomics or Master of Human Factors degrees.

More recently, occupational therapists have been moving into the field of ergonomics and the field has been heralded as one of the top ten emerging practice areas to watch for in the new millennium in occupational therapy.

Applications

The more than twenty technical subgroups within the Human Factors and Ergonomics Society, HFES, indicate the range of applications for ergonomics. Human factors engineering continues to be successfully applied in the fields of aerospace, aging, health care, IT, product design, transportation, training, nuclear and virtual environments, among others. Kim Vicente, a University of Toronto Professor of Ergonomics, argues that the nuclear disaster in Chernobyl is attributable to plant designers not paying enough attention to human factors. "The operators were trained but the complexity of the reactor and the control panels nevertheless outstripped their ability to grasp what they were seeing [during the prelude to the disaster]."

Physical ergonomics is important in the medical field, particularly to those diagnosed with physiological ailments or disorders such as arthritis (both chronic and temporary) or carpal tunnel syndrome. Pressure that is insignificant or imperceptible to those unaffected by these disorders may be very painful, or render a device unusable, for those who are. Many ergonomically designed products are also used or recommended to treat or prevent such disorders, and to treat pressure-related chronic pain.

Human factors issues arise in simple systems and consumer products as well. Some examples include cellular telephones and other handheld devices that continue to shrink yet grow more complex (a phenomenon referred to as "creeping featurism"), millions of VCRs blinking "12:00" across the world because very few people can figure out how to program them, or alarm clocks that allow sleepy users to inadvertently turn off the alarm when they mean to hit 'snooze'. A user-centered design (UCD), also known as a systems approach or the usability engineering lifecycle aims to improve the user-system

Engineering Psychology

Engineering psychology is an interdisciplinary part of Ergonomics and studies the relationships of people to machines, with the intent of improving such relationships. This may involve redesigning equipment, changing the way people use machines, or changing the location in which the work takes place. Often, the work of an engineering psychologist is described as making the relationship more "user-friendly."

Engineering Psychology is an applied field of psychology concerned with psychological factors in the design and use of equipment. Human factors is broader than engineering psychology, which is focused specifically on designing systems that accommodate the information-processing capabilities of the brain (see Wickens and Hollands 2000).

Cognitive Load

Cognitive Load is a term (used in psychology and other fields of study) that refers to the load on working memory during problem solving, thinking and reasoning (including perception, memory, language, etc.).

Most would agree that people learn better when they can build on what they already understand. But the more things a person has to learn in a short amount of time, the more difficult it is to process information in working memory.

Consider the difference between having to study a subject in one's native language versus trying to study a subject in a foreign language. The cognitive load is much higher in the second instance because the brain must work to translate the language while simultaneously trying to understand the new information.

Another aspect of cognitive load theory involves understanding how many discrete units of information can be retained in short term memory before information loss occurs. An example that seems to be commonly cited of this principle is the use of 7-digit phone numbers, based on the theory that most people can only retain seven "chunks" of information in their short term memory. Refer to Chunking (psychology).

Cognitive Load Theory, as defined by Sweller (1988), states that optimum learning occurs in humans when the load on working memory is kept to a minimum to best facilitate the changes in long term memory.

The History Of Cognitive Load Theory

The history of cognitive load theory can be traced back to the beginning of Cognitive Science and the work of G.A. Miller (1956). Miller was perhaps the first to suggest our working memory capacity was limited in his classic paper (Miller, 1956). He suggested we are only able to hold seven plus or minus two digits of information in our short term memories. Miller's early work was built upon by many researchers in the ensuing decades. Perhaps most notably by that of Simon and Chase (1973) who also used the term "chunk" to describe how experts use their short term memories. As novices learn, they begin to see patterns in the world around them. These patterns can be combined with other patterns... this chunking of memory components has also been described as schema construction.

John Sweller developed cognitive load theory while studying problem solving (Sweller, 1988). While studying learners as they solved problems, he and his associates found that learners often use a problem solving strategy called means-ends analysis. He suggests problem solving by means-ends analysis requires a relatively large amount of cognitive processing capacity, which may not be devoted to schema construction. Instead of problem solving, Sweller suggests Instructional designers should limit cognitive load by designing instructional materials like worked-examples, or goal-free problems.

In the 1990s, Cognitive load theory was applied in several contexts and the empirical results from these studies led to the demonstration of several learning effects: the completion-problem effect (Paas, 1992); Modality effect (Moreno & Mayer, 1999; Mousavi, Low, & Sweller, 1995); Split-attention effect (Chandler and Sweller, 1992); and the Worked-example effect (Sweller & Cooper, 1985; Cooper & Sweller, 1987).

The Human Cognitive Architecture & Instructional Design

Cognitive load theory has been used to describe the architecture of human cognition. It has been suggested that Cognitive load has broad implications for Instructional design (Sweller, 1999). This theory provides a general framework for instructional designers for it allows them to control the conditions of learning within an environment or more generally within most instructional materials. Specifically it provides empirically-based guidelines that help instructional designers to minimize extraneous cognitive load during learning.

John Sweller's theory employs information processing theory to emphasize the inherent limitations of working memory. In addition it uses schemas as the relevant unit of analysis for the design of instructional materials.

This theory differentiates between three types of cognitive load: intrinsic cognitive load, germane cognitive load, and extraneous cognitive load (Sweller, van Merriënboer, and Paas, 1998).

Intrinsic Cognitive Load

The term "Intrinsic cognitive load" was first described by Chandler and Sweller (1991). Accordingly all instruction has an inherent difficulty associated with it (e.g., the calculation of 2 + 2, versus solving a differential equation). This inherent difficulty may not be altered by an instructor. However many schemas may be broken into individual "subschemas" and taught in isolation, to be later brought back together and described as a combined whole (Clark, Nguyen, and Sweller, 2006).

Extraneous Cognitive Load

Extraneous load is that load which instructional designers do have some ability to control. This load can be attributed to the design of the instructional materials.

Sweller provides a wonderful example of extraneous cognitive load in his 2006 book, when he describes two possible ways to describe a square to a student (Clark, Nguyen, and Sweller, 2006). A square is a visual and should be described using a visual medium. Certainly an instructor can describe a square in a verbal medium, but it takes just a second and far less effort to see what the instructor is talking about when a learner is shown a square, rather than having one described verbally. In this instance, the efficiency of the visual medium is preferred. This is because it does not unduly load the learner with unnecessary information. This unnecessary cognitive load is described as extraneous cognitive load.

Germane Cognitive Load

Germane load was first described by Sweller, van Merrienboer and Paas in 1998. It is that load devoted to the processing, construction and automation of schemata. While intrinsic load is generally thought to be immutable, instructional designers can manipulate extraneous and germane load. It is suggested that they limit extraneous load and promote germane load (Sweller, van Merrienboer & Paas, 1998).

Human Factors

"Human Factors" are sets of human-specific physical, mental and behavioral properties which either may interact in a critical or dangereous manner with technological systems, human natural environment and human organizations, or they can be taken under consideration in the design of ergonomic human-user oriented equipments. The choice/identification of human factors usually depends on their possible negative or positive impact on the functioning of human-organization and human-machine system.

"Human Factors" also is the name of an engineering profession that focuses on how people interact with tasks, machines (or computers), and the environment with the consideration that humans have limitations and capabilities. Often, human factors will study the human within the system to ensure that we understand the limitations of the human within the current structure, product, or process. Human factors engineers will evaluate human to human, human to group, or human to organizational interactions to better understand the phenomena associated with these interactions and to develop a framework for evaluation.

The term "human factors" is used mainly in the United States. Variants include "human factors engineering", an extension of an earlier phrase, and "human engineering". In the rest of the world, the term "ergonomics" is more prevalent. Cognitive ergonomics is another term used. Within the US, the term "ergonomics" tends to apply to the physical and physiological phenomena for optimal human interactions, whereas "human factors" is rather used to refer to a focus on psychological and psychosocial factors in systems(e.g., behavioral, cognitive, perceptual, psychodynamic).

Human factors practitioners can come from a variety of backgrounds; though predominantly they are Psychologists (Cognitive, Perceptual, and Experimental) and industrial engineers. Designers (Industrial, Interaction, and Graphic), Anthropologists, Technical communication Scholars and Computer Scientists also contribute. Though some practitioners enter the field of Human Factors from other disciplines, both M.S. and Ph.D. degrees in Human Factors Engineering are available from several universities worldwide, and degreed professionals can obtain professional certification from the Board of Certified Professional Ergonomists(BCPE).

Areas of interest for human factors practitioners may include: training, staffing evaluation, communication, task analyses, functional requirements analyses and allocation, job descriptions and functions, procedures and procedure use, knowledge, skills, and abilities; organizational culture, human-machine interaction, workload on the human, fatigue, situational awareness, usability, user interface, learnability, attention, vigilance, human performance, human reliability, human-computer interaction, control and display design, stress, visualization of data, individual differences, aging, accessibility, safety, shift work, work in extreme environments including virtual environments, human error, and decision making.

Simply put, human factors involves working to make the environment function in a way that seems natural to people and attempts to optimize tasks, the machine design, and/or the environment. Although the terms "human factors" and "ergonomics" have only been widely known in recent times, the field's origin is in the design and use of aircraft during World War II to improve aviation safety. It was in reference to the psychologists working at that time and the work that they were doing that the terms "applied psychology" and "applied psychologist" were first coined.

The Human-Machine Model

The simple human-machine model is of a person interacting with a machine in some kind of environment. The person and machine are both modeled as information-processing devices, each with inputs, central processing, and outputs. The inputs of a person are the senses (e.g., eyes, ears) and the outputs are effectors (e.g., hands, voice). The inputs of a machine are input control devices (e.g., keyboard, mouse) and the outputs are output display devices (e.g., screen, auditory alerts).

The environment can be characterized physically (e.g., vibration, noise, zero-gravity), cognitively (e.g., time pressure, uncertainty, risk), and/or organizationally (e.g., organizational structure, job design). This provides a convenient way for organizing some of the major concerns of human engineering:

- the selection and design of machine displays and controls;
- the layout and design of workplaces;
- design for maintainability;
- and the design of the work environment.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited*****.io***

Other - The Unbanked

Who Are They?

The unbanked are described by the Federal Deposit Insurance Corporation (FDIC) as those without an account at a bank or other financial institution and are considered to be outside the mainstream for one reason or another. The FDIC estimates there are 10 million unbanked or underbanked American households. The majority of them are American born while a growing number are immigrants where the two groups have low income as a commonality and lack the minimum balance to open checking and savings accounts. According to Congressman Hinojosa, half of the unbanked had a bank account previously but are choosing to not have an account and opting to using the services of check cashers and payday lenders instead.

In 2003, Donald E. Powell, Chairman of the FDIC, mentioned a personal experience that caused him to believe that banks need to make a greater effort to ensure "all are welcome in the lobbies" which includes getting out into the community because trust is important.

Governor Arnold Schwarzenegger started the Bank on California initiative to help the unbanked in 2008. Previously, in 2001, a financial education curriculum called Money Smart was launched by the FDIC to help the financially unsavvy.

Prior to becoming the FDIC chair in 2006, Sheila Bair ran a research project at for the Inter-American Development Bank at the University of Massachusetts to discovery ways to help unbanked Latin American immigrants use the U.S. Banking System. Like past studies it was found that the primary reason recent Latino immigrants don't use banks to remit money is because they don't have documents of legal alien status. Around the same time the Treasury Department put in place Section 326 regulations that allow banks and credit unions to accept identification from foreign governments at their own discretion. Banks like Mitchell Bank in Milwaukee have taken up the Treasury Department on their relaxing of identification standards. They have even "offered pamphlets on how to apply for a Wisconsin state ID and driver's license, and invited the Mexican consulate in Chicago to visit with a mobile unit that issues "matricula" cards." In Chicago, the Consul General of Mexico, Carlos Sada, estimated that up to 25% of applicants of the Matricula Consular ID apply in order use it to acquire U.S. bank accounts.

Other - Development Economics

What Is It?

Development economics is a branch of economics which deals with economic aspects of the development process in low-income countries. Its focus is not only on methods of promoting economic growth and structural change but also on improving the potential for the mass of the population, for example, through health and education and workplace conditions, whether through public or private channels. Thus, development economics involves the creation of theories and methods that aid in the determination of types of policies and practices and can be implemented at either the domestic or international level. This may involve restructuring market incentives or using mathematical methods like inter-temporal optimization for project analysis, or it may involve a mixture of quantitative and qualitative methods. Unlike in many other fields of economics, approaches in development economics may incorporate social and political factors to devise particular plans. Different approaches may consider the factors that contribute to economic convergence or non-convergence across households, regions, and countries.

Theories Of Development Economics

Mercantilism

The earliest modern Western theory of development economics was mercantilism, which developed in the 17th century, paralleling the rise of the nation state. Under the earlier system of medieval feudalism, Scholasticism was the dominant school of thought, and emphasized reconciliation with Christian theology and ethics, rather than development. The succeeding 16th and 17th century School of Salamanca, credited as the earliest modern school of economics, likewise did not address development specifically.

Major European nations in the 17th and 18th century all adopted mercantilist ideals to varying degrees, the influence only ebbing with the 18th century development of physiocrats in France and classical economics in Britain. Mercantilist development theory also advocated colonialism.

Theorists most associated with mercantilism include Philipp Wilhelm von Hornick, who in his Austria Over All, If She Only Will of 1684 gave the only comprehensive statement of mercantilist theory, emphasizing production and an export-led economy. In France, mercantilist policy is most associated with 17th century finance minister Jean-Baptiste Colbert, whose policies proved influential in later American development.

Mercantilist ideas continue in the theories of economic nationalism and neomercantilism.

Economic Nationalism

Following mercantilism was the related theory of economic nationalism, promulgated in the 19th century related to the development and industrialization of the United States and Germany, notably in the policies of the American System in America and the Zollverein (customs union) in Germany. A significant difference from mercantilism was the deemphasis on colonies, in favor of a focus on domestic production.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

The names most associated with 19th century economic nationalism are the American Alexander Hamilton, the German-American Friedrich List, and the American Henry Clay. Hamilton's 1791 Report on Manufactures, his magnum opus, is the founding text of the American System, and drew from the mercantilist economies of Britain under Elizabeth I and France under Colbert. List's 1841 Das Nationale System der Politischen Ökonomie (translated into English as The National System of Political Economy), which emphasized stages of growth, proved influential in the US and Germany, and nationalist policies were pursued by politician Henry Clay, and later by Abraham Lincoln, under the influence of economist Henry Charles Carey.

Forms of economic nationalism and neomercantilism have also been key in Japan's development in the 19th and 20th centuries, and the more recent development of the Four Asian Tigers (Hong Kong, South Korea, Taiwan, and Singapore), and, most significantly, China.

Post-WWII Theories

The origins of modern development economics are often traced to the need for, and likely problems with the industrialization of Eastern Europe in the aftermath of World War II. The key authors are Paul Rosenstein-Rodan, Kurt Mandelbaum, Ragnar Nurkse, and Sir Hans Wolfgang Singer. Only after the war did economists turn their concerns towards Asia, Africa and Latin America. At the heart of these studies, by authors such as Simon Kuznets and W. Arthur Lewis was an analysis of not only economic growth but also structural transformation.

Linear-Stages-Of-Growth Model

An early theory of development economics, the linear-stages-of-growth model was first formulated in the 1950s by W. W. Rostow in The Stages of Growth: A Non-Communist Manifesto, following work of Marx and List. This theory modifies Marx's stages theory of development and focuses on the accelerated accumulation of capital, through the utilization of both domestic and international savings as a means of spurring investment, as the primary means of promoting economic growth and, thus, development. The linear-stages-of-growth model posits that there are a series of five consecutive stages of development which all countries must go through during the process of development. These stages are “the traditional society, the pre-conditions for take-off, the take-off, the drive to maturity, and the age of high mass-consumption” Simple versions of the Harrod–Domar model provide a mathematical illustration of the argument that improved capital investment leads to greater economic growth.

Such theories have been criticized for not recognizing that, while necessary, capital accumulation is not a sufficient condition for development. That is to say that this early and simplistic theory failed to account for political, social and institutional obstacles to development. Furthermore, this theory was developed in the early years of the Cold War and was largely derived from the successes of the Marshall Plan. This has led to the major criticism that the theory assumes that the conditions found in developing countries are the same as those found in post-WWII Europe.

Structural-Change Theory

Structural-change theory deals with policies focused on changing the economic structures of developing countries from being composed primarily of subsistence agricultural practices to being a “more modern, more urbanized, and more industrially diverse manufacturing and service economy.” There are two major forms of structural-change theory; W. Lewis’ two-sector surplus model, which views agrarian societies as consisting of large amounts of surplus labor which can be utilized to spur the development of an urbanized industrial sector, and Hollis Chenery's patterns of development approach, which is the empirical analysis of the "sequential process through which the economic, industrial and institutional structure of an underdeveloped economy is transformed over time to permit new industries to replace traditional agriculture as the engine of economic growth."

Structural-change approaches to development economics have faced criticism for their emphasis on urban development at the expense of rural development which can lead to a substantial rise in inequality between internal regions of a country. The two-sector surplus model, which was developed in the 1950s, has been further criticized for its underlying assumption that predominantly agrarian societies suffer from a surplus of labor. Actual empirical studies have shown that such labor surpluses are only seasonal and drawing such labor to urban areas can result in a collapse of the agricultural sector. The patterns of development approach have been criticized for lacking a theoretical framework.

International Dependence Theory

International dependence theories gained prominence in the 1970s as a reaction to the failure of earlier theories to lead to widespread successes in international development. Unlike earlier theories, international dependence theories have their origins in developing countries and view obstacles to development as being primarily external in nature, rather than internal. These theories view developing countries as being economically and politically dependent on more powerful, developed countries which have an interest in maintaining their dominant position. There are three different, major formulations of international dependence theory; neocolonial dependence theory, the false-paradigm model and the dualistic-dependence model. The first formulation of international dependence theory, neocolonial dependence theory has its origins in Marxism and views the failure of many developing nations to undergo successful development as being the result of the historical development of the international capitalist system.

Neoclassical Theory

First gaining prominence with the rise of several conservative governments in the West during the 1980s, neoclassical theories represent a radical shift away from International Dependence Theories. Neoclassical theories argue that governments should not intervene in the economy; in other words, these theories are claiming that an unobstructed free market is the best means of inducing rapid and successful development. Competitive free markets unrestrained by excessive government regulation are seen as being able to naturally ensure that the allocation of resources occurs with the greatest efficiency possible and the economic growth is raised and stabilized.

It is important to note that there are several different approaches within the realm of neoclassical theory, each with subtle, but important, differences in their views regarding the extent to which the market should be left unregulated. These different takes on neoclassical theory are the free market approach, public-choice theory, and the market-friendly approach. Of the three, both the free-market approach and public-choice theory contend that the market should be totally free, meaning that any intervention by the government is necessarily bad. Public-choice theory is arguably the more radical of the two with its view, closely associated with libertarianism, that governments themselves are rarely good and therefore should be as minimal as possible.

The market-friendly approach, unlike the other two, is a more recent development and is often associated with the World Bank. This approach still advocates free markets but recognizes that there are many imperfections in the markets of many developing nations and thus argues that some government intervention is an effective means of fixing such imperfections.

Topics Of Research

Development economics also includes topics such as Third World debt, and the functions of such organisations as the International Monetary Fund and World Bank. In fact, the majority of development economists are employed by, do consulting with, or receive funding from institutions like the IMF and the World Bank. Many such economists are interested in ways of promoting stable and sustainable growth in poor countries and areas, by promoting domestic self reliance and education in some of the lowest income countries in the world. Where economic issues merge with social and political ones, it is referred to as development studies.

Growth Indicator Controversy

Per capita Gross Domestic Product (GDP per head) is used by many developmental economists as an approximation of general national well-being. However, these measures are criticized as not measuring economic growth well enough, especially in countries where there is much economic activity that is not part of measured financial transactions (such as housekeeping and self-homebuilding), or where funding is not available for accurate measurements to be made publicly available for other economists to use in their studies (including private and institutional fraud, in some countries). Even though per-capita GDP as measured can make economic well-being appear smaller than it really is in some developing countries, the discrepancy could be still bigger in a developed country where people may perform outside of financial transactions an even higher-value service than housekeeping or homebuilding as gifts or in their own households, such as counseling, lifestyle coaching, a more valuable home décor service, and time management. Even free choice can be considered to add value to lifestyles without necessarily increasing the financial transaction amounts. More recent theories of Human Development have begun to see beyond purely financial measures of development, for example with measures such as medical care available, education, equality, and political freedom. One measure used is the Genuine Progress Indicator, which relates strongly to theories of distributive justice. Actual knowledge about what creates growth is largely unproven; however recent advances in econometrics and more accurate measurements in many countries is creating new knowledge by compensating for the effects of variables to determine probable causes out of merely correlational statistics.

Recent Developments

The most prominent contemporary development economist is perhaps the Nobel laureate Amartya Sen. Recent theories revolve around questions about what variables or inputs correlate or affect economic growth the most: elementary, secondary, or higher education, government policy stability, tariffs and subsidies, fair court systems, available infrastructure, availability of medical care, prenatal care and clean water, ease of entry and exit into trade, and equality of income distribution (for example, as indicated by the Gini coefficient), and how to advise governments about macroeconomic policies, which include all policies that affect the economy. Education enables countries to adapt the latest technology and creates an environment for new innovations. The cause of limited growth and divergence in economic growth lies in the high rate of acceleration of technological change by a small number of developed countries. These countries' acceleration of technology was due to increased incentive structures for mass education which in turn created a framework for the population to create and adapt new innovations and methods. Furthermore, the content of their education was composed of secular schooling that resulted in higher productivity levels and modern economic growth.

Prominent Development Economists

- Paul Collier, author of The Bottom Billion which attempts to tie together a series of traps to explain the self-fulfilling nature of poverty at the lower end of the development scale.
- Ha-Joon Chang, author of Kicking Away the Ladder and Bad Samaritans; Rich Nations, Poor Policies and the Threat to the Developing World which use historical evidence to critique neoliberal development economics.
- Joseph Stiglitz, Nobel Prize winner and former chief economist at the World Bank.
- Dani Rodrik
- Lant Pritchett
- Kaushik Basu, professor of the economics at Cornell University and author of Analytical Development Economics.
- William Easterly, author of The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics and White Man's Burden: How the West's Efforts to Aid the Rest have done so much ill and so little good
- Arthur Lewis
- Patrick Lynch, author of Economic development and planning with Basil Chubb, Readings in Irish public Administration; Volume 1, Institute of Public Administration, Dublin, 1969.
- Raúl Prebisch, founding Secretary General of the United Nations Conference on Trade and Development and influential dependency theorist
- Walt Whitman Rostow, modernization theorist, author of The Stages of Economic Growth: A Non-communist Manifesto
- Jeffrey Sachs, author of The End of Poverty: Economic Possibilities of Our Time and Common Wealth: Economics for a Crowded Planet
- Amartya Sen, Nobel Prize winner, author of Development as Freedom
- Hans Singer, who dealt with how unequal terms of trade disproportionately affect producers of primary products. His thesis, combined with the work of Raúl Prebisch, form the basis for dependency theory
- Hernando de Soto Polar, proponent of property rights in the developing world, author of The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else
- Frances Stewart
- Celso Furtado
- Esther Duflo
- Abhijit Banerjee

Other - Inclusive Business

What Is It?

An inclusive business is a sustainable business that benefits low-income communities.

Large corporations traditionally target consumers in the middle and high-income segments of society, and established suppliers and service providers from the formal economy. Inclusive businesses find profitable ways to engage the low-income segment into their business operations in a way that benefits the low-income communities and creates sustainable livelihoods.

Inclusive businesses may engage low-income communities through, among other things, directly employing low-income people; targeting development of suppliers and service providers from low-income communities; or providing affordable goods and services targeted at low-income communities.

Inclusive business is not corporate philanthropy, which has inherent limitations of scope and budget. Rather, it is the search for sustainable business models that "do well by doing good" and have the potential to become part of the mainstream business model within the companies concerned - the key to business having development impact at scale.

As employees and suppliers, the low-income segments gain access to the formal economy; including provision of training, access to finance and income. As consumers, low income customers can benefit from products and services that meet their needs in an affordable way. If business does both, it opens up the virtuous cycle of business in development.

Other - Emerging Markets

What Is It?

The term emerging markets is used to describe a nation's social or business activity in the process of rapid growth and industrialization. Currently, there are approximately 28 emerging markets in the world, with the economies of India and China considered to be by far the two largest. According to The Economist many people find the term dated, but a new term has yet to gain much traction.

Terminology

Originally brought into fashion in the 1980s by then World Bank economist Antoine van Agtmael, the term is sometimes loosely used as a replacement for emerging economies, but really signifies a business phenomenon that is not fully described by or constrained to geography or economic strength; such countries are considered to be in a transitional phase between developing and developed status. Examples of emerging markets include Argentina, Brazil, Chile, China, Colombia, India, Mexico, Peru, much of Southeast Asia, countries in Eastern Europe and in the Middle East, and parts of Africa and Latin America. Emphasizing the fluid nature of the category, political scientist Ian Bremmer defines an emerging market as "a country where politics matters at least as much as economics to the markets."

The research on emerging markets is diffused within management literature. While researchers including C. K. Prahalad, George Haley, Hernando de Soto, Usha Haley, and several professors from Harvard Business School and Yale School of Management have described activity in countries such as India and China, how a market emerges is little understood.

In the 2008 Emerging Economy Report the Center for Knowledge Societies defines Emerging Economies as those "regions of the world that are experiencing rapid informationalization under conditions of limited or partial industrialization." It appears that emerging markets lie at the intersection of non-traditional user behavior, the rise of new user groups and community adoption of products and services, and innovations in product technologies and platforms.

The term "rapidly developing economies" is being used to denote emerging markets such as The United Arab Emirates, Chile and Malaysia that are undergoing rapid growth.

In recent years, new terms have emerged to describe the largest developing countries such as BRIC that stand for Brazil, China, India and Russia. These countries do not share any common agenda, but some experts believe that they are enjoying an increasing role in the world economy and on political platforms.

A large number of research works are in progress at leading universities and business schools to study and understand various aspects of Emerging Markets.

It is difficult to make an exact list of emerging (or developed) markets; the best guides tend to be investment information sources like ISI Emerging Markets and The Economist or market index makers (such as Morgan Stanley Capital International). These sources are well-informed, but the nature of investment information sources leads to two potential problems. One is an element of historicity; markets may be maintained in an index for continuity, even if the countries have since developed past the emerging market phase. Possible examples of this are Israel, South Korea, and Taiwan. A second is the simplification inherent in making an index; small countries, or countries with limited market liquidity are often not considered, with their larger neighbours considered an appropriate stand-in.

The Big Emerging Market (BEM) economies are Brazil, China, Egypt, India, Indonesia, Mexico, Philippines, Poland, Russia, South Africa, South Korea and Turkey.[10]

Newly industrialized countries are emerging markets whose economies have not yet reached first world status but have, in a macroeconomic sense, outpaced their developing counterparts.

FTSE Emerging Markets List

The FTSE Group distinguishes between Advanced and Secondary Emerging Markets on the basis of their national income and the development of their market infrastructure. The Advanced Emerging Markets are classified as such because they are Upper Middle Income GNI countries with advanced market infrastructures or High Income GNI countries with lesser developed market infrastructures.

The Advanced Emerging Markets are: Brazil, Hungary, Mexico, Poland, South Africa, Taiwan.

The Secondary Emerging Markets are some Upper Middle, Lower Middle and Low Income GNI countries with reasonable market infrastructures and significant size and some Upper Middle Income GNI countries with lesser developed market infrastructures.

The Secondary Emerging Markets are: Argentina, Chile, China, Colombia, Czech Republic, Egypt, India, Indonesia, Malaysia, Morocco, Pakistan, Peru, Philippines, Russia, Thailand, Turkey.

MSCI List

As of April 2009, MSCI Barra classified the following 22 countries as emerging markets:

- Brazil
- Chile
- China
- Colombia
- Czech Republic
- Egypt
- Hungary
- India
- Indonesia
- Israel
- Malaysia
- Mexico
- Morocco
- Peru
- Philippines
- Poland
- Russia
- South Africa
- South Korea
- Taiwan

- Thailand
- Turkey

The list tracked by The Economist is the same, except with Hong Kong, Singapore and Saudi Arabia included (MSCI classifies the first two as Developed Markets)

Other - Recession

What Is It?

In economics, a recession is a general slowdown in economic activity over a long period of time, or a business cycle contraction.[1][2] During recessions, many macroeconomic indicators vary in a similar way. Production as measured by Gross Domestic Product (GDP), employment, investment spending, capacity utilization, household incomes and business profits all fall during recessions.

Governments usually respond to recessions by adopting expansionary macroeconomic policies, such as increasing money supply, increasing government spending and decreasing taxation.

Identifying

In a 1975 New York Times article, economic statistician Julius Shiskin suggested several rules of thumb for identifying a recession, one of which was "two down quarters of GDP". In time, the other rules of thumb were forgotten, and a recession is now often defined simply as a period when GDP falls (negative real economic growth) for at least two quarters. Some economists prefer a definition of a 1.5% rise in unemployment within 12 months.

In the United States the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER) is generally seen as the authority for dating US recessions. The NBER defines an economic recession as: "a significant decline in [the] economic activity spread across the country, lasting more than a few months, normally visible in real GDP growth, real personal income, employment (non-farm payrolls), industrial production, and wholesale-retail sales." Almost universally, academics, economists, policy makers, and businesses defer to the determination by the NBER for the precise dating of a recession's onset and end.

Attributes

A recession has many attributes that can occur simultaneously and includes declines in coincident measures of activity such as employment, investment, and corporate profits.

A severe (GDP down by 10%) or prolonged (three or four years) recession is referred to as an economic depression, although some argue that their causes and cures can be different. As an informal shorthand, economists sometimes refer to different recession shapes, such as V-shaped, U-shaped, L-shaped and W-shaped recessions.

In the US, V-shaped, or short-and-sharp contractions followed by rapid and sustained recovery, occurred in 1954 and 1990-91; U-shaped (prolonged slump) in 1974-75, and W-shaped, or double-dip recessions in 1949 and 1980-82. Japan’s 1993-94 recession was U-shaped and its 8-out-of-9 quarters of contraction in 1997-99 can be described as L-shaped. Korea, Hong Kong and South-east Asia experienced U-shaped recessions in 1997-98, although Thailand’s eight consecutive quarters of decline should be termed L-shaped.

Government Responses

Most mainstream economists believe that recessions are caused by inadequate aggregate demand in the economy, and favor the use of expansionary macroeconomic policy during recessions. Strategies favored for moving an economy out of a recession vary depending on which economic school the policymakers follow. Monetarists would favor the use of expansionary monetary policy, while Keynesian economists may advocate increased government spending to spark economic growth. Supply-side economists may suggest tax cuts to promote business capital investment. Laissez-faire minded economists may simply recommend that the government not interfere with natural market forces.

Other - Slum

What Is It?

A slum, as defined by the United Nations agency UN-HABITAT, is a run-down area of a city characterized by substandard housing and squalor and lacking in tenure security. According to the United Nations, the proportion of urban dwellers living in slums decreased from 47 percent to 37 percent in the developing world between 1990 and 2005. However, due to rising population, the number of slum dwellers is rising. One billion people worldwide live in slums and will likely grow to 2 billion by 2030.

The term has traditionally referred to housing areas that were once respectable but which deteriorated as the original dwellers moved on to newer and better parts of the city, but has come to include the vast informal settlements found in cities in the developing world.

Many shack dwellers vigorously oppose the description of their communities as 'slums' arguing that this results in them being pathologised and then, often, subject to threats of evictions. Many academics have vigorously criticized UN-Habitat and the World Bank arguing that their 'Cities Without Slums' Campaign has led directly to a massive increase in forced evictions.

Although their characteristics vary between geographic regions, they are usually inhabited by the very poor or socially disadvantaged. Slum buildings vary from simple shacks to permanent and well-maintained structures. Most slums lack clean water, electricity, sanitation and other basic services.

Characteristics

The characteristics associated with slums vary from place to place. Slums are usually characterized by urban decay, high rates of poverty, and unemployment. They are commonly seen as "breeding grounds" for social problems such as crime, drug addiction, alcoholism, high rates of mental illness, and suicide. In many poor countries they exhibit high rates of disease due to unsanitary conditions, malnutrition, and lack of basic health care.

A UN Expert Group has created an operational definition of a slum as an area that combines to various extents the following characteristics: inadequate access to safe water; inadequate access to sanitation and other infrastructure; poor structural quality of housing; overcrowding; and insecure residential status. A more complete definition of these can be found in the 2003 UN report titled "Slums of the World: The face of urban poverty in the new millennium?". The report also lists various attributes and names that are given by individual countries which are somewhat different than these UN characteristics of a slum.

Low socioeconomic status of its residents is another common characteristic given for a slum.

In many slums, especially in poor countries, many live in very narrow alleys that do not allow vehicles (like ambulances and fire trucks) to pass. The lack of services such as routine garbage collection allows rubbish to accumulate in huge quantities. The lack of infrastructure is caused by the informal nature of settlement and no planning for the poor by government officials. Additionally, informal settlements often face the brunt of natural and man-made disasters, such as landslides, as well as earthquakes and tropical storms. Fires are often a serious problem.

Many slum dwellers employ themselves in the informal economy. This can include street vending, drug dealing, domestic work, and prostitution. In some slums people even recycle trash of different kinds (from household garbage to electronics) for a living - selling either the odd usable goods or stripping broken goods for parts or raw materials.

Slums are often associated with Victorian Britain, particularly in industrial, northern towns. These were generally still inhabited until the 1940s, when the government started slum clearance and built new council houses. There are still many examples left of former slum housing in the UK, however they have generally been restored into more modern housing.

Growth And Countermeasures

Recent years have seen a dramatic growth in the number of slums as urban populations have increased in the Third World.

In April 2005, the director of UN-HABITAT stated that the global community was falling short of the Millennium Development Goals which targeted significant improvements for slum dwellers and an additional 50 million people have been added to the slums of the world in the past two years. According to a 2006 UN-HABITAT report, 327 million people live in slums in Commonwealth countries - almost one in six Commonwealth citizens. In a quarter of Commonwealth countries (11 African, 2 Asian and 1 Pacific), more than two out of three urban dwellers live in slums and many of these countries are urbanising rapidly.

The number of people living in slums in India has more than doubled in the past two decades and now exceeds the entire population of Britain, the Indian Government has announced.

Many governments around the world have attempted to solve the problems of slums by clearing away old decrepit housing and replacing it with modern housing with much better sanitation. The displacement of slums is aided by the fact that many are squatter settlements whose property rights are not recognized by the state. This process is especially common in the Third World. Slum clearance often takes the form of eminent domain and urban renewal projects, and often the former residents are not welcome in the renewed housing. Moreover new projects are often on the semi-rural peripheries of cities far from opportunities for generating livelihoods as well as schools, clinics etc. At times this has resulted in large movements of inner city slum dwellers militantly opposing relocation to formal housing on the outskirts of cities. For example Abahlali baseMjondolo in Durban, South Africa.

In some countries, leaders have addressed this situation by rescuing rural property rights to support traditional sustainable agriculture, however this solution has met with open hostility from capitalists and corporations. It also tends to be relatively unpopular with the slum communities themselves, as it involves moving out of the city back into the countryside, a reverse of the rural-urban migration that originally brought many of them into the city.

Critics argue that slum clearances tend to ignore the social problems that cause slums and simply redistribute poverty to less valuable real estate. Where communities have been moved out of slum areas to newer housing, social cohesion may be lost. If the original community is moved back into newer housing after it has been built in the same location, residents of the new housing face the same problems of poverty and powerlessness. There is a growing movement to demand a global ban of 'slum clearance programmes' and other forms of mass evictions.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited.**io

Other – Social Network

What Is It?

A social network is a social structure made of nodes (which are generally individuals or organizations) that are tied by one or more specific types of relations, such as values, visions, idea, financial exchange, friends, kinship, dislike, trade, web links, sexual relations, disease transmission (epidemiology), or airline routes.

Social network analysis views social relationships in terms of nodes and ties. Nodes are the individual actors within the networks, and ties are the relationships between the actors. There can be many kinds of ties between the nodes. Research in a number of academic fields has shown that social networks operate on many levels, from families up to the level of nations, and play a critical role in determining the way problems are solved, organizations are run, and the degree to which individuals succeed in achieving their goals.

In its simplest form, a social network is a map of all of the relevant ties between the nodes being studied. The network can also be used to determine the social capital of individual actors. These concepts are often displayed in a social network diagram, where nodes are the points and ties are the lines.

Social Network Analysis

Social network analysis (related to network theory) has emerged as a key technique in modern sociology, anthropology, sociolinguistics, geography, social psychology, communication studies, information science, organizational studies, economics, and biology as well as a popular topic of speculation and study.

People have used the social network metaphor for over a century to connote complex sets of relationships between members of social systems at all scales, from interpersonal to international. Yet not until J. A. Barnes in 1954 did social scientists start using the term systematically to denote patterns of ties that cut across the concepts traditionally used by the public and social scientists: bounded groups (e.g., tribes, families) and social categories (e.g., gender, ethnicity). Social network analysis developed with the kinship studies of Elizabeth Bott in England in the 1950s and the urbanization studies "Manchester School" (centered around Max Gluckman and later J. Clyde Mitchell), done mainly in Zambia during the 1960s. It joined with the field of sociometry, begun by Jacob L. Moreno in the 1930s as an attempt to quantify social relationships. Scholars such as S.D. Berkowitz, Stephen Borgatti, Ronald Burt, Linton Freeman, Mark Granovetter, Nicholas Mullins, Stanley Wasserman, Barry Wellman and Harrison White expanded the use of social networks.

Social network analysis has now moved from being a suggestive metaphor to an analytic approach to a paradigm, with its own theoretical statements, methods and research tribes. Analysts reason from whole to part; from structure to relation to individual; from behavior to attitude. They either study whole networks, all of the ties containing specified relations in a defined population, or personal networks, the ties that specified people have, such as their "personal communities".

Several analytic tendencies distinguish social network analysis:

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited*****.io***

- There is no assumption that groups are the building blocks of society: the approach is open to studying less-bounded social systems, from nonlocal communities to links among Web sites.
- Rather than treating individuals (persons, organizations, states) as discrete units of analysis, it focuses on how the structure of ties affects individuals and their relationships.
- By contrast with analyses that assume that socialization into norms determines behaviour, network analysis looks to see the extent to which the structure and composition of ties affect norms.

The shape of a social network helps determine a network's usefulness to its individuals. Smaller, tighter networks can be less useful to their members than networks with lots of loose connections (weak ties) to individuals outside the main network. More open networks, with many weak ties and social connections, are more likely to introduce new ideas and opportunities to their members than closed networks with many redundant ties. In other words, a group of friends who only do things with each other already share the same knowledge and opportunities. A group of individuals with connections to other social worlds is likely to have access to a wider range of information. It is better for individual success to have connections to a variety of networks rather than many connections within a single network. Similarly, individuals can exercise influence or act as brokers within their social networks by bridging two networks that are not directly linked (called filling structural holes).

The power of social network analysis stems from its difference from traditional social scientific studies, which assume that it is the attributes of individual actors - whether they are friendly or unfriendly, smart or dumb, etc. - that matter. Social network analysis produces an alternate view, where the attributes of individuals are less important than their relationships and ties with other actors within the network. This approach has turned out to be useful for explaining many real-world phenomena, but leaves less room for individual agency, the ability for individuals to influence their success, because so much of it rests within the structure of their network.

Social networks have also been used to examine how organizations interact with each other, characterizing the many informal connections that link executives together, as well as associations and connections between individual employees at different organizations. For example, power within organizations often comes more from the degree to which an individual within a network is at the center of many relationships than actual job title. Social networks also play a key role in hiring, in business success, and in job performance. Networks provide ways for companies to gather information, deter competition, and even collude in setting prices or policies.

History Of Social Network Analysis

A comprehensive summary of the theoretical progress of social networks and social network analysis is written by Linton Freeman. His book, The development of social network analysis, gives great insight in the development of the social theories.

Precursors of social networks in the late 1800s include Emile Durkheim and Ferdinand Tönnies. Tönnies argued that social groups can exist as personal and direct social ties that either link individuals who share values and belief ("gemeinschaft") or impersonal, formal and instrumental social links ("gesellschaft"). Durkheim gave a non-individualistic explanation of social facts arguing that social phenomena arise when interacting individuals constitute a reality that can no longer be accounted for in terms of the properties of individual actors (Coser, 1977). He distinguished between a traditional society – "mechanical solidarity" – which prevails if individual differences are minimalized, and the modern society – "organic solidarity" – which develops out of differences of individuals what results in cooperation of individuals.

Georg Simmel, writing at the turn of the twentieth century, was the first scholar to think directly in social network terms. His essays pointed to the nature of network size on interaction and to the likelihood of interaction in ramified, loosely-knit networks rather than groups (Simmel, 1908/1971).

After a hiatus in the first decades of the twentieth century, three main traditions in social networks appeared. In the 1930s, one tradition worked on sociometric analysis of small groups with Kohler and J.L. Moreno, and a Harvard group with W. Lloyd Warner and Elton Mayo who explored interpersonal relationships and ties. In the 1950s-1960s, anthropologists centered around the University of Manchester, such as J. Clyde Mitchell and Elizabeth Bott investigated community networks in southern Africa, India and the United Kingdom.

In the 1960’s, Harrison White at Harvard University was able to combine the different tracks and traditions (see Scott, 2000 and Freeman, 2004 for elaborate explanations of all three traditions and White’s work). Mark Granovetter and Barry Wellman are former students of White who have elaborated and popularized social networks.

Applications

The evolution of social networks can sometimes be modeled by the use of agent based models, providing insight into the interplay between communication rules, rumor spreading and social structure. Here is an interactive model of rumour spreading, based on rumour spreading from model on Cmol.

Diffusion of innovations theory explores social networks and their role in influencing the spread of new ideas and practices. Change agents and opinion leaders often play major roles in spurring the adoption of innovations, although factors inherent to the innovations also play a role.

Dunbar's number: The so-called rule of 150, asserts that the size of a genuine social network is limited to about 150 members (sometimes called Dunbar's number). The rule arises from cross-cultural studies in sociology and especially anthropology of the maximum size of a village (in modern parlance most reasonably understood as an ecovillage). It is theorized in evolutionary psychology that the number may be some kind of limit of average human ability to recognize members and track emotional facts about all members of a group. However, it may be due to economics and the need to track "free riders", as it may be easier in larger groups to take advantage of the benefits of living in a community without contributing to those benefits.

Guanxi is a central concept in Chinese society that can be summarized as the use of personal influence. Guanxi can be studied from a social network approach.

The small world phenomenon is the hypothesis that the chain of social acquaintances required to connect one arbitrary person to another arbitrary person anywhere in the world is generally short. The concept gave rise to the famous phrase six degrees of separation after a 1967 small world experiment by psychologist Stanley Milgram. In Milgram's experiment, a sample of US individuals were asked to reach a particular target person by passing a message along a chain of acquaintances. The average length of successful chains turned out to be about five intermediaries or six separation steps (the majority of chains in that study actually failed to complete). Academic researchers continue to explore this phenomenon. Judith Kleinfeld has written an article that points out the many problems with the original Milgram research. A recent electronic Small World experiment at Columbia University showed that about five to seven degrees of separation are sufficient for connecting any two people through e-mail.

The study of socio-technical systems is loosely linked to social network analysis, and looks at relations among individuals, institutions, objects and technologies.

Metrics (Measures) In Social Network Analysis

- Betweenness - Degree an individual lies between other individuals in the network; the extent to which a node is directly connected only to those other nodes that are not directly connected to each other; an intermediary; liaisons; bridges. Therefore, it's the number of people who a person is connected to indirectly through their direct links.

- Centrality Closeness - The degree an individual is near all other individuals in a network (directly or indirectly). It reflects the ability to access information through the "grapevine" of network members. Thus, closeness is the inverse of the sum of the shortest distances between each individual and every other person in the network.

- Centrality Degree - The count of the number of ties to other actors in the network.

- Flow Betweenness Centrality - The degree that a node contributes to sum of maximum flow between all pairs of nodes (not that node).

- Centrality Eigenvector - Eigenvector centrality is a measure of the importance of a node in a network. It assigns relative scores to all nodes in the network based on the principle that connections to nodes having a high score contribute more to the score of the node in question.

- Centralization - The difference between the n of links for each node divided by maximum possible sum of differences. A centralized network will have much of its links dispersed around one or a few nodes, while a decentralized network is one in which there is little variation between the n of links each node possesses

- Clustering Coefficient - The clustering coefficient is a measure of the likelihood that two associates of a node are associates themselves. A higher clustering coefficient indicates a greater 'cliquishness'.

- Cohesion - Refers to the degree to which actors are connected directly to each other by cohesive bonds. Groups are identified as ‘cliques’ if every actor is directly tied to every other actor, ‘social circles’ if there is less stringency of direct contact, which is imprecise, or as structurally cohesive blocks if precision is wanted.

- Contagion

- Density - Individual-level density is the degree a respondent's ties know one another/ proportion of ties among an individual's nominees. Network or global-level density is the proportion of ties in a network relative to the total number possible (sparse versus dense networks).

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited*****.io***

- Integration
- Path Length - The distances between pairs of nodes in the network. Average path-length is the average of these distances between all pairs of nodes.
- Radiality - Degree an individual's network reaches out into the network and provides novel information and influence
- Reach - The degree any member of a network can reach other members of the network.
- Structural Cohesion - The minimum number of members who, if removed from a group, would disconnect the group.
- Structural Equivalence - Refers to the extent to which actors have a common set of linkages to other actors in the system. The actors don't need to have any ties to each other to be structurally equivalent.
- Structural Hole - Static holes that can be strategically filled by connecting one or more links to link together other points. Linked to ideas of social capital: if you link to two people who are not linked you can control their communication.

Other – Social Network Service

What Is It?

A social network service focuses on the building and verifying of online social networks for communities of people who share interests and activities, or who are interested in exploring the interests and activities of others, and which necessitates the use of software.

Most social network services are primarily web based and provide a collection of various ways for users to interact, such as chat, messaging, email, video, voice chat, file sharing, blogging, discussion groups, and so on.

There have been some attempts to standardize them but this has led to some privacy concerns.

Business Applications

Social networks connect people at low cost; this can be beneficial for entrepreneurs and small businesses looking to expand their contact base. These networks often act as a customer relationship management tool for companies selling products and services. Companies can also use social networks for advertising in the form of banners and text ads. Since businesses operate globally, social networks can make it easier to keep in touch with contacts around the world.

Medical Applications

Social networks are beginning to be adopted by healthcare professionals as a means to manage institutional knowledge, disseminate peer-to-peer knowledge and to highlight individual physicians and institutions. The advantage of using a dedicated medical social networking site is that all the members are screened against the state licensing board list of practitioners.

The role of social networks is especially of interest to pharmaceutical companies who spend approximately "32 percent of their marketing dollars" attempting to influence the opinion leaders of social networks.

Typical Structure Of A Social Networking Service

Basics

In general, social networking services, such as MySpace, Facebook and Bebo, allow users to create a profile for themselves. Users can upload a picture of themselves and can often be "friends" with other users. In most social networking services, both users must confirm that they are friends before they are linked. For example, if Alice lists Bob as a friend, then Bob would have to approve Alice's friend request before they are listed as friends. Some social networking sites have a "favourites" feature that does not need approval from the other user. Social networks usually have privacy controls that allows the user to choose who can view their profile or contact them, etc.

Additional Features

Some social networks have additional features, such as the ability to create groups that share common interests or affiliations, upload videos, and hold discussions in forums.

Business Networking Sites

Business networking sites have absorbed the traditional face-to-face referral networking practices of realtors and title companies, attorneys and chiropractors, and businesses networkers from a variety of fields, and taken it online. Businesses from all across the globe can come together and share ideas, clients, and referrals in a mutually beneficial manner. Such sites allow users to network with one another online, without ever having to meet someone face-to-face. Some of these services are entirely free.

Device-Based Social Networking

In addition to internet-based social networking, cell phone manufacturers are getting into the social networking business with phones that allow users to create lists of friends and associates, track their movements even across countries, and create customized maps and alerts that signal the user when a desired person is within a predetermined range. With such GPS-enabled phones, users are able to send out invitations or messages to groups of people based on customized attributes, including location.

User Behaviour

Users often try to "collect friends", or try to be linked to as many friends as possible. Therefore, it is not uncommon for users to receive friend requests from people that they do not know. Some users will create additional profiles that assume the identity of someone else, such as celebrities, politicians, or even their pets. Some will create profiles for fictional characters, such as those from video games or films (similar to role-playing), and some will even create profiles for inanimate objects, such as the Sun or the dwarf planet Pluto.

Privacy Issues

On large social networking services, there have been growing concerns about users giving out too much personal information and the threat of sexual predators. Users of these services need to be aware of data theft or viruses. However, large services, such as MySpace, often work with law enforcement to try to prevent such incidents.

In addition, there is a perceived privacy threat in relation to placing too much personal information in the hands of large corporations or governmental bodies, allowing a profile to be produced on an individual's behaviour on which decisions, detrimental to an individual, may be taken.

Furthermore, there is an issue over the control of data - information having been altered or removed by the user may in fact be retained and/or passed to 3rd parties.

Investigations

Social network services are increasingly being used in legal and criminal investigations. Information posted on sites such as MySpace and Facebook, has been used by police and university officials to prosecute users of said sites. In some situations, content posted on MySpace has been used in court to determine an appropriate sentence based on a defendant's attitude.

Facebook is increasingly being used by school administrations and law enforcement agencies as a source of evidence against student users. The site, the number one online destination for college students, allows users to create profile pages with personal details. These pages can be viewed by other registered users from the same school which often include resident assistants and campus police who have signed-up for the service.

Other – Six Degrees Of Separation

What Is It?

Six degrees of separation refers to the idea that, if a person is one "step" away from each person he or she knows and two "steps" away from each person who is known by one of the people he or she knows, then everyone is no more than six "steps" away from each person on Earth. Several studies, such as Milgram's small world experiment, have been conducted to empirically measure this connectedness. While the exact number of links between people differs depending on the population measured, it is generally found to be relatively small. Hence, six degrees of separation is somewhat synonymous with the idea of the "small world" phenomenon.

Early Conceptions

The "Shrinking World"

In 1929, a Hungarian author named Frigyes Karinthy published a volume of short stories titled "Everything is Different." One of these pieces was titled "Chains," or "Chain-Links." The story investigated in abstract, conceptual, and fictional terms many of the problems that would captivate future generations of mathematicians, sociologists, and physicists within the field of network theory. In particular, Karinthy believed that the modern world was shrinking due to the ever-increasing connectedness of human beings. Due to technological advances in communications and travel, friendship networks could grow larger and span even greater distances. Karinthy posited that despite great physical distances between the globe's individuals, the growing density of human networks made the actual social distance far smaller.

As a result of this hypothesis, Karinthy's characters believed that any two individuals could be connected through at most five acquaintances. In his story, the characters create a game out of this notion. He writes:

> A fascinating game grew out of this discussion. One of us suggested performing the following experiment to prove that the population of the Earth is closer together now than they have ever been before. We should select any person from the 1.5 billion inhabitants of the Earth—anyone, anywhere at all. He bet us that, using no more than five individuals, one of whom is a personal acquaintance, he could contact the selected individual using nothing except the network of personal acquaintances.

This idea both directly and indirectly influenced a great deal of early thought on social networks. Thus, Karinthy is often regarded as the originator of the notion of Six Degrees of Separation.

The "Small World" Experiments

Stanley Milgram was an American researcher in experimental social psychology at Harvard University in Boston, USA. Beginning in 1967, he began a widely-publicized set of experiments to investigate the so-called "small world problem." This problem was rooted in

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.planetlimited.io

many of the same observations made decades earlier by Karinthy. That is, Milgram and other researchers of the era were fascinated by the interconnectedness and "social capital" of human networks. While it is unknown how directly Milgram was influenced by Karinthy's work, the similarities between the two authors are remarkable. However, while Karinthy spoke in abstract and fictional terms, Milgram's experiments provided evidence supporting the claim of a "small world." His study results showed that people in the United States seemed to be connected by approximately six friendship links, on average. Although Milgram reportedly never used the term "Six Degrees of Separation," his findings likely contributed to the term's widespread credence. Since these studies were widely publicized, Stanley Milgram is also, like Karinthy, often attributed as the origin of the notion of Six Degrees.

Theoretical Basis & Proof

If you assume the world population is 6 billion people, and everyone has the same amount of friends or 'connections', and that each person is just as likely to know to know one person as any oher person (save for geographic limitations), then measuring the degree of separation only becomes a simple mathematical formula of determining the exponent that will yield the population if you raise the average number of friends of each person by that exponent. In other words,

(average number of friends per person) ^ (degrees of separation) = total population
Let f = average number of friends
Let d = degress of separation
Let p = population
f^d = p
*d * ln f = ln p*
d = ln p / ln f

Finding the average number of friends can be determined by random sampling. However, since we already have a good idea what the degree of freedom is, let's determine 'f' and consider its reasonableness.

f^d = p
f = dth root of p

If we take the 6th root of 6,000,000,000, we get approximately 42.

6th root of 6,000,000,000 = 42.628
ln(6,000,000,000) / ln(42) = 6.024

Knowing 42 people is not unreasonable to a person. One class or workplace can contain 42 people. Thus, if 42 of your friends knows 42 other people, and they each know 42 more people, and so on and so on until 6 chains have been formed, then that will encompass 6 billion people.

In fact, if we take the 5th root of 6 billion, we get about 90, which in today's connected age is not unreasonable for some people. The 7th root is 24. So if we assume that everyone in the world knows between 24 to 90 people each, then we can prove the degree of separation is between 5 and 7.

The reason this works is because as the number of chains increase, the total percentage of the population 'known' increases exponentially. For example, if the population is 16, and each person is restricted to knowing at most 2 people, then the degree of separation is 4. 1 person who knows 2 people is 2. If those 2 people know 2 more people, the total is 4. If those 4 people know 2 people each, the total is 8. If those 8 people know 2 people each, the total is 16.

4th root of 16 = 2
2^4 = 16

If the population is split in two groups of 8, perhaps by geographic boundries, then it would be impossible to 'know' the entire population or have a connection between all individuals.

The advent of affordable intercontinental air travel in the 20th century has reduced these geographic boundaries, such that even if one individual does not know someone on another continent, they are likely to know someone else who has been to another continent. The more the population intermixes and comingles, the more even and regular the degree of separation is between any two random people in that population.

Recent Research

In 2001, Duncan Watts, a professor at Columbia University, attempted to recreate Milgram's experiment on the internet, using an e-mail message as the "package" that needed to be delivered, with 48,000 senders and 19 targets (in 157 countries). Watts found that the average (though not maximum) number of intermediaries was around six. This finding is surprising, given the worldwide nature of the Internet.

It has been suggested by some commentators that interlocking networks of computer mediated lateral communication could diffuse single messages to all interested users worldwide as per the 6 degrees of separation principle via Information Routing Groups, which are networks specifically designed to exploit this principle and lateral diffusion.

SixDegrees.org

The game "Six Degrees of Kevin Bacon" was invented in 1994 by two students at Albright College in Reading, Pennsylvania, as a play on the concept: the goal is to link any actor to Kevin Bacon through no more than six connections, where two actors are connected if they have appeared in a movie together.

On January 18, 2007, Kevin Bacon launched SixDegrees.org, a web site that builds on the popularity of the "small world phenomenon" to create a charitable social network and inspire giving to charities online. Bacon started the network with celebrities who are highlighting their favorite charities – including Kyra Sedgwick (Natural Resources Defense Council), Nicole Kidman (UNIFEM), Ashley Judd (YouthAIDS), Bradley Whitford and Jane Kaczmarek (Clothes off Our Back), Dana Delany (Scleroderma Research Foundation), Robert Duvall (Pro Mujer), Rosie O'Donnell (Rosie's For All Kids Foundation), and Jessica Simpson (Operation Smile) - and he encouraged everyone to be celebrities for their own causes by joining the Six Degrees movement.

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.**planetlimited*****.io***

"SixDegrees.org is about using the idea that we are all connected to accomplish something good," said Bacon. "It is my hope that Six Degrees will soon be something more than a game or a gimmick. It will also be a force for good, by bringing a social conscience to social networking." The game, 'Six Degrees of Kevin Bacon,' made the rounds of college campuses over the past decade and lived on to be a shorthand term for the small world phenomenon.

Bacon created SixDegrees.org in partnership with the non-profit, Network for good, AOL, and Entertainment Weekly. Through SixDegrees.org, which builds on Network for Good's giving system for donating to more than one million charities online and AOL's AIM Pages social networking service, people can learn about and support the charities of celebrities or fundraise for their own favourite causes with their own friends and families. Bacon will match the charitable dollars raised by the top six non-celebrity fundraisers with grants of up to $10,000 each.

Other – Push-Pull Strategy

What Is It?

Push-Pull Strategy

The business terms push and pull originated in the marketing and advertising world, but are also applicable in the world of electronic content and supply chain management. The push/pull relationship is that between a product or piece of information and who is moving it. A customer "pulls" things towards themselves, while a producer "pushes" things toward customers.

Overview

In a "push" system the consumer does not request the product to be developed; it is "pushed at" the end-user by promotion. An example of this is a perfume product. Women do not request to smell a fragrance they never smelled before; it is simply "pushed" at them, through the right advertisement.

- Applied to that portion of the supply chain where demand uncertainty is relatively small
- Production & distribution decisions are based on long term forecasts
- Based on past orders received from retailer’s warehouse (may lead to the Bullwhip effect)
- Inability to meet changing demand patterns
- Large and variable production batches
- Unacceptable service levels
- Excessive inventories due to the need for large safety stocks

In a "pull" system the consumer requests the product and "pulls" it through the delivery channel. An example of this is the car manufacturing company Toyota. Toyota only produces cars when they have been ordered by the customers.

- Applied to that portion of the supply chain where demand uncertainty is high
- Production and distribution are demand driven
- No inventory, response to specific orders
- Point of sale (POS) data comes in handy when shared with supply chain partners
- Decrease in lead time
- Difficult to implement

Supply Chains

With a push-based supply chain, products are pushed through the channel, from the production side up to the retailer. The manufacturer sets production at a level in accord with historical ordering patterns from retailers. It takes longer for a push-based supply chain to respond to changes in demand, which can result in overstocking or bottlenecks and delays (the bullwhip effect), unacceptable service levels and product obsolescence.

In a pull-based supply chain, production and distribution are demand driven so that they are coordinated with actual customer orders, rather than forecasted demand.

A supply chain is almost always a combination of both push and pull, where the interface between the push-based stages and the pull-based stages is known as the push-pull boundary. An example of this would be Dell's build to order supply chain. Inventory levels of individual components are determined by forecasting general demand, but final assembly is in response to a specific customer request. The push-pull boundary would then be at the beginning of the assembly line. At this point on the supply chain timeline, it is typically coordinated through a buffer inventory.

The Bullwhip Effect

The Bullwhip Effect (or Whiplash Effect) is an observed phenomenon in forecast-driven distribution channels. The concept has its roots in J Forrester's Industrial Dynamics (1961). Because customer demand is rarely perfectly stable, businesses must forecast demand in order to properly position inventory and other resources. Forecasts are based on statistics, and they are rarely perfectly accurate. Because forecast errors are a given, companies often carry an inventory buffer called "safety stock". Moving up the supply chain from end-consumer to raw materials supplier, each supply chain participant has greater observed variation in demand and thus greater need for safety stock. In periods of rising demand, down-stream participants will increase their orders. In periods of falling demand, orders will fall or stop in order to reduce inventory. The effect is that variations are amplified as one moves upstream in the supply chain (further from the customer).

Bullwhip effect is also attributed to the separate ownership of different stages of the supply chain. Each stage in such a structured supply chain tries to amplify the profit of the respective stages, thereby decreasing the overall profitability of the supply chain.

Supply chain experts have recognized that the Bullwhip Effect is a problem in forecast-driven supply chains, and careful management of the effect is an important goal for Supply Chain Managers. The alternative is to establish a demand-driven supply chain which reacts to actual customer orders. In manufacturing, this concept is called Kanban. This model has been most successfully implemented in Wal-Mart's distribution system. Individual Wal-Mart stores transmit point-of-sale (POS) data from the cash register back to corporate headquarters several times a day. This demand information is used to queue shipments from the Wal-Mart distribution center to the store and from the supplier to the Wal-Mart distribution center. The result is near-perfect visibility of customer demand and inventory movement throughout the supply chain. Better information leads to better inventory positioning and lower costs throughout the supply chain. Barriers to implementing a demand-driven supply chain include investments in information technology and creating a corporate culture of flexibility and focus on customer demand.

Factors contributing to the Bullwhip Effect:

- Forecast Errors
- Lead Time Variability
- Batch Ordering
- Price Fluctuations
- Product Promotions
- Inflated Orders

Methods intended to reduce uncertainty, variability, and lead time:

- Vendor Managed Inventory (VMI)
- Just In Time replenishment (JIT)

- Strategic partnerships

Other – Digital Rights Management

What Is It?

Digital rights management (DRM) is an umbrella term that refers to access control technologies used by publishers and copyright holders to limit usage of digital media or devices. It may also refer to restrictions associated with specific instances of digital works or devices. DRM overlaps with software copy protection to some extent, however the term "DRM" is usually applied to creative media (music, films, etc.) whereas the term "copy protection" tends to refer to copy protection mechanisms in computer software.

Digital rights management has and is being used by content provider companies such as Sony, Apple Inc., Microsoft and the BBC.

The use of digital rights management has been controversial. Advocates argue it is necessary for copyright holders to prevent unauthorized duplication of their work to ensure continued revenue streams. Opponents, such as the Free Software Foundation, maintain that the use of the word "rights" is misleading and suggest that people instead use the term digital restrictions management. Their position is essentially that copyright holders are attempting to restrict use of copyrighted material in ways not covered by existing laws. The Electronic Frontier Foundation, and other opponents, also consider DRM systems to be anti-competitive practices.

In practice, all widely-used DRM systems have been defeated or circumvented when deployed to enough customers. Protection of audio and visual material is especially difficult due to the existence of the analog hole, and there are even suggestions that effective DRM is logically impossible for this reason.

Orientation

Digital rights management technologies attempt to control use of digital media by preventing access, copying or conversion by end users to other formats. Long before the arrival of digital or even electronic media, copyright holders, content producers, or other financially or artistically interested parties had business and legal objections to copying technologies. Examples include: player piano rolls early in the 20th century, audio tape recording, and video tape recording (e.g. the "Betamax case" in the U.S.). Copying technology thus exemplifies a disruptive technology.

The advent of digital media and analog/digital conversion technologies, especially those that are usable on mass-market general-purpose personal computers, have vastly increased the concerns of copyright-dependent organizations, especially within the music and movie industries. While analog media inevitably loses quality with each copy generation, and in some cases even during normal use, digital media files may be duplicated an unlimited number of times with no degradation in the quality of subsequent copies. The advent of personal computers as household appliances has made it convenient for consumers to convert media (which may or may not be copyrighted) originally in a physical/analog form or a broadcast form into a universal, digital form (this process is called ripping) for location and/or time shifting purposes, combined with the Internet and popular file sharing tools, has made unauthorized distribution of copies of copyrighted digital media (so-called digital piracy) much easier. In effect, copyright-dependent organizations regard every consumer with an Internet connection as potential nodes in a distribution network that could be used to distribute unauthorized copies of copyrighted works.

Although technical controls on the reproduction and use of software have been intermittently used since the 1970s, the term 'DRM' has come to primarily mean the use of these measures to control artistic or literary content. DRM technologies have enabled publishers to enforce access policies that not only disallow copyright infringements, but also prevent lawful fair use of copyrighted works, or even implement use constraints on non-copyrighted works that they distribute; examples include the placement of DRM on certain public-domain or open-licensed e-books, or DRM included in consumer electronic devices that time-shift (and apply DRM to) both copyrighted and non-copyrighted works.

While DRM is most commonly used by the entertainment industry (e.g. film and recording), it has found use in other situations as well. Many online music stores, such as Apple's iTunes Store, as well as certain e-book publishers, have imposed DRM on their customers. In recent years, a number of television producers have imposed DRM mandates on consumer electronic devices, to control access to the freely-broadcast content of their shows, in connection with the popularity of time-shifting digital video recorder systems such as TiVo.

Technologies - DRM & Film

An early example of a DRM system was the Content Scrambling System (CSS) employed by the DVD Forum on film DVDs since circa 1996. CSS used a simple encryption algorithm, and required device manufacturers to sign license agreements that restricted the inclusion of features, such as digital outputs that could be used to extract high-quality digital copies of the film, in their players. Thus, the only consumer hardware capable of decoding DVD films was controlled, albeit indirectly, by the DVD Forum, restricting the use of DVD media on other systems until the release of DeCSS by Jon Lech Johansen in 1999, which allowed a CSS-encrypted DVD to play properly on a computer using Linux, for which the Alliance had not arranged a licensed version of the CSS playing software.

Microsoft's Windows Vista contains a DRM system called the Protected Media Path, which contains the Protected Video Path (PVP). PVP tries to stop DRM-restricted content from playing while unsigned software is running in order to prevent the unsigned software from accessing the content. Additionally, PVP can encrypt information during transmission to the monitor or the graphics card, which makes it more difficult to make unauthorized recordings.

Advanced Access Content System (AACS) is a DRM system for HD DVD and Blu-Ray Discs developed by the AACS Licensing Administrator, LLC (AACS LA), a consortium that includes Disney, Intel, Microsoft, Matsushita (Panasonic), Warner Brothers, IBM, Toshiba and Sony. In December 2006 a process key was published on the internet by crackers, enabling unrestricted access to AACS-restricted HD DVD content. After the cracked keys were revoked, further cracked keys were released.

The broadcast flag concept was developed by Fox Broadcasting in 2001 and was supported by the MPAA and the FCC. A ruling in May 2005 by a US Court of Appeals held that the FCC lacked authority to impose it on the TV industry in the US. It required that all HDTVs obey a stream specification determining whether or not a the stream can be recorded. This could block instances of fair use, such as time-shifting. It achieved more success elsewhere when it was adopted by the Digital Video Broadcasting Project (DVB), a consortium of about 250 broadcasters, manufactures, network operators, software developers, and regulatory bodies from about 35 countries involved in attempting to develop new digital TV standards.

An updated variant of the broadcast flag has been developed in the Content Protection and

Copy Management (DVB-CPCM). It was developed in private, and the technical specification was submitted to European in March 2007. As with much DRM, the CPCM system is intended to control use of copyrighted material by the end-user, at the direction of the copyright holder. According to Ren Bucholz of the EFF, which paid to be a member of the consortium, "You won't even know ahead of time whether and how you will be able to record and make use of particular programs or devices". The DVB supports the system as it will harmonize copyright holders control across different technologies and so make things easier for end users. The CPCM system is expected to be submitted to the European Telecommunications Standards Institute in 2008.

In March 2008, the Canadian Broadcasting Corporation has announced that it would release a DRM-free version of one of its television programs called Canada's Next Great Prime Minister through BitTorrent.

Technologies - DRM & Music

Audio CDs

It should be noted that discs with DRM installed are not legitimately standards-compliant Compact Discs (CDs) but rather CD-ROM media, therefore they all lack the CD logotype found on discs which follow the standard (known as Red Book). Therefore these CDs could not be played on all CD players. Many consumers could also no longer play purchased CDs on their computers. PCs running Microsoft Windows would sometimes even crash when attempting to play the CDs.

In 2002, Bertelsmann (comprising BMG, Arista, and RCA) was the first corporation to use DRM on audio CDs. In 2005, Sony BMG introduced new DRM technology which installed DRM software on user's computers, without clearly notifying the user or requiring their confirmation. Among other things, the installed software included a rootkit, which created a severe security vulnerability others could exploit. When the nature of the DRM involved was made public much later, Sony initially minimized the significance of the vulnerabilities its software had created, but was eventually compelled to recall millions of CDs, and released several attempts to patch the surreptitiously included software to at least remove the rootkit. Several class action lawsuits were filed, which were ultimately settled by agreements to provide affected consumers with a cash payout or album downloads free of DRM.

Sony's DRM software actually had only a limited ability to prevent copying, as it affected only playback on Windows computers, not on other equipment. Even on the Windows platform, users regularly bypassed the restrictions. And, while the Sony DRM technology created fundamental vulnerabilities in customers' computers, parts of it could be trivially bypassed by holding down the "shift" key while inserting the CD, or by disabling the autorun feature. In addition, audio tracks could simply be played and re-recorded, thus completely bypassing all of the DRM (this is known as the analog hole). Sony's first two attempts at releasing a patch which would remove the DRM software from users' computers failed.

In January 2007, EMI stopped publishing audio CDs with DRM, stating that "the costs of DRM do not measure up to the results." EMI was the last publisher to do so, and audio CDs containing DRM are no longer released by any major publishers.

Internet Music

Many online music stores employ DRM to restrict usage of music purchased and downloaded online. There are many options for consumers buying digital music over the internet, in terms of both stores and purchase options.

- The iTunes Store, run by Apple Inc., allows users to purchase a track online for $.99 US. The tracks purchased use Apple's FairPlay DRM system. Starting on October 17, 2007, users can download DRM-free music for the same price as a file with DRM.
- Napster music store, which offers a subscription-based approach to DRM alongside permanent purchases. Users of the subscription service can download and stream an unlimited amount of music encoded to Windows Media Audio (WMA) while subscribed to the service. But as soon as the user misses a payment, the service renders all of the downloaded music unusable. Napster also charges users who wish to use the music on their portable device an additional $5 per month. Furthermore, Napster requires users to pay an additional $0.99 per track to burn it to CD or listen to it after the subscription expires. Songs bought through Napster can be played on players carrying the Microsoft PlaysForSure logo (which, notably, do not include iPods or even Microsoft's own Zune).
- Wal-Mart Music Downloads, another online music download store, charges $0.94 per track for all non-sale downloads. All Wal-Mart Music Downloads are able to be played on any Windows PlaysForSure marked product. The music does play on the SanDisk's Sansa mp3 player, for example, but must be copied to the player's internal memory. It can not be played through the player's microSD card slot, which is a problem that many users of the mp3 player experience.
- Sony operated an online music download service called "Connect" which used Sony's proprietary OpenMG DRM technology. Music downloaded from this store (usually via Sony's SonicStage software) was only playable on computers running Windows and Sony hardware (including the PSP).

The various services are currently not interoperable, though those that use the same DRM system (for instance the several Windows Media DRM format stores, including Napster) all provide songs that can be played side-by-side through the same player program. Almost all stores require client software of some sort to be downloaded, and some also need plug-ins. Several colleges and universities, such as Rensselaer Polytechnic Institute, have made arrangements with assorted Internet music suppliers to provide access (typically DRM-restricted) to music files for their students, to less than universal popularity, sometimes making payments from student activity fee funds. One of the problems is that the music becomes unplayable after leaving school unless the student continues to pay individually. Another is that few of these vendors are compatible with the most common portable music player, the Apple iPod. The Gowers Review of Intellectual Property (to HMG in the UK; 141 pages, 40+ specific recommendations) has taken note of the incompatibilities, and suggests (Recommendations 8 -- 12) that there be explicit fair dealing exceptions to copyright allowing libraries to copy and format-shift between DRM schemes, and further allowing end users to do the same privately. If adopted, some of the acrimony may decrease.

Although DRM is prevalent for Internet music, some Online music stores such as eMusic, Audio Lunchbox, Dogmazic, Amazon and Anthology recordings do not use DRM. Major labels have begun releasing more online music without DRM. Eric Bangeman suggests in Ars Technica that this is because the record labels are "slowly beginning to realize that they can't have DRMed music and complete control over the online music market at the same time... One way to break the cycle is to sell music that is playable on any digital audio player. eMusic does exactly that, and their surprisingly extensive catalog of non-DRMed music has vaulted it into the number two online music store position behind the iTunes Store." Apple's Steve Jobs has called on the music industry to eliminate DRM in an open letter titled Thoughts on Music. Apple's iTunes store will start to sell DRM-free 256 kbit/s (up from 128 kbit/s) music from EMI for a premium price (this has since reverted to the standard price). In March of 2007, Musicload.de, one of Europe's largest online music retailers, announced their position strongly against DRM. In an open letter, Musicload stated that three out of every four calls to their customer support phone service are as a result of consumer frustration with DRM.

E-Books

Electronic books on the PC typically use DRM restrictions, but some systems, such as eReader (formerly Palm Digital Media) instead write the purchaser’s credit card info into the title.

Technologies - DRM & Documents

Enterprise digital rights management (E-DRM or ERM) is the application of DRM technology to the control of access to corporate documents such as Microsoft Word, PDF, and AutoCAD files, emails, and intranet web pages rather than to the control of consumer media. E-DRM, now more commonly referenced as IRM (Information Rights Management), is generally intended to prevent the unauthorized use (such as industrial or corporate espionage or inadvertent release) of proprietary documents. IRM typically integrates with content management system software.

DRM has been used by organizations such as the British Library in its secure electronic delivery service to permit worldwide access to substantial numbers of rare (and in many cases unique) documents which, for legal reasons, were previously only available to authorized individuals actually visiting the Library's document centre at Boston Spa in England.

Watermarks

Digital watermarks are unobtrusive features of media that are added during production or distribution. Digital watermarks involve data steganographically embedded within the audio or video data.

Watermarks can be used for different purposes that may include:

- for recording the copyright owner
- for recording th
- e distributor
- for recording the distribution chain
- for identifying the purchaser of the music

Watermarks are not complete DRM mechanisms in their own right, but are used as part of a system for Digital Rights Management, such as helping provide prosecution evidence for purely legal avenues of rights management, rather than direct technological restriction.

Metadata

Sometimes, metadata is included in purchased music which records information such as the purchaser's name, account information, or email address. This information is not embedded in the played audio or video data, like a watermark, but is kept separate, but within the file or stream.

As an example, metadata is used in media purchased from Apple's iTunes Store for DRM-free as well as DRM-restricted versions of their music or videos. This information is included as MPEG standard metadata.

Laws Regarding DRM

Digital rights management systems have received some international legal backing by implementation of the 1996 WIPO Copyright Treaty (WCT). Article 11 of the Treaty requires nations party to the treaties to enact laws against DRM circumvention.

The WCT has been implemented in most member states of the World Intellectual Property Organization. The American implementation is the Digital Millennium Copyright Act (DMCA), while in Europe the treaty has been implemented by the 2001 European directive on copyright, which requires member states of the European Union to implement legal protections for technological prevention measures. In 2006, the lower house of the French parliament adopted such legislation as part of the controversial DADVSI law, but added that protected DRM techniques should be made interoperable, a move which caused widespread controversy in the United States.

Controversy

The intent of DRM is to provide technical means to assure that the copyright holders (originally artists, but commonly assigned to publishers, software developers, etc.) can maintain control of their content by restricting use of digital copies. This becomes controversial because DRM imposed limitations on the use of legitimately acquired digital media do not necessarily match the fair use (fair dealing in some places) rights granted by law to owners of copies. This gives rise to concerns that DRM schemes enormously complicate, and may prevent, effective archive management and historical research as well. Others argue that DRM is ineffective at preventing illegal copies because no DRM technology is (or could possibly be) fool proof. Once one version is compromised (or simply copied from a medium without DRM) it will become widely available, e.g. on the Internet or via large-scale commercial piracy. Thus all DRM to date is claimed to be fundamentally technically flawed as a method of protecting legal copyright control. If so, its effect is essentially to ensure vendor lock-in and, likely, anti-competitive practices afterwards. DRM opponents usually base their opposition on one or more of these concerns.

Additional arguments against DRM are based on the fact that Copyright Laws limit the duration of copyrights, requiring that the DRM-restricted material be placed into the public domain at the end of the granted copyright period. DRM systems violate this requirement of copyright law in as much as DRM systems are not programmed to terminate at the end of the copyright period, effectively extending the "copyright" beyond what is allowable by law. As such, this use of DRM is arguably itself a violation of the same copyright law that the proponents of DRM claim the system enforces.

Shortcomings Of DRM

Methods To Bypass DRM

There are many methods to bypass DRM control on audio and video content.

One simple method to bypass DRM on audio files is to burn the content to an audio CD and then rip it into DRM-free files. This is only possible when the software that plays these DRM-restricted audio files allows CD-burning. Some software products simplify and automate this burn-rip process by allowing the user to burn music to a CD-RW disc or to a Virtual CD-R drive, then automatically ripping and encoding the music, and automatically repeating this process until all selected music has been converted, rather than forcing the user to do this one CD (72-80 minutes' worth of music) at a time.

Many software programs have been developed that intercept the data stream as it is decrypted out of the DRM-restricted file, and then use this data to construct a DRM-free file. These programs require a decryption key. Programs that do this for DVDs, HD DVDs, and Blu-ray Discs include universal decryption keys in the software itself. Programs that do this for TiVo ToGo recordings, iTunes audio, and PlaysForSure songs, however, rely on the user's own key, meaning they can only process content that the user has legally acquired under his or her own account.

Another method is to use software to record the signals being sent through the audio or video cards, or to plug analog recording devices into the analog outputs of the media player. These techniques utilize the so-called "analog hole".

Analog Hole

All forms of DRM for audio and visual material are subject to the analog hole, namely that in order for a viewer to play the material, the digital signal must be turned into an analog signal containing light and/or sound for the viewer, and so available to be copied as no DRM is capable of controlling content in this form. In other words, a user could playback a purchased audio file while using a separate program to record the sound back into the computer into a non-DRM protected file format.

All DRM to date, and probably all future ones can therefore be bypassed by recording this signal and digitally storing and distributing it in a non DRM limited form. However the conversion from digital to analog and back is likely to force a loss of quality, particularly when using lossy digital formats. HDCP is an attempt to restrict the analog hole.

DRM On General Computing Platforms

Many of the DRM systems in use are designed to work on general purpose computing hardware, such as desktop PCs apparently because this equipment is felt to be a major contributor to revenue loss from disallowed copying. Large commercial pirates avoid consumer equipment, so losses from such infringers will not be covered by such provisions.

Some have suggested that any such scheme can never be secure since the software must include all the information, such as decryption keys, necessary to decrypt the content. It is suggested that one can always extract this information and decrypt and copy the content, bypassing the restrictions imposed by a DRM system.

DRM On Distributed Purpose Built Hardware

Many DRM schemes use encrypted media which requires purpose built hardware to hear or see the content. This appears to ensure that only licensed users (those with the hardware) can access the content. It additionally tries to protect a secret decryption key from the users of the system.

While this in principle can work, it is extremely difficult to build the hardware to protect the secret key against a sufficiently determined adversary. Many such systems have failed in the field, and in fact, it is thought that none have yet survived several years of deployment. Once the secret key is known, building a version of the hardware that performs no checks is often relatively straightforward.

In addition user verification provisions are frequently subject to attack.

Watermarks

Watermarks are usually easily removed, although some degradation of video or audio can occur.

In particular, most compression is intended to only retain perceptible features of an image, and hence if the watermarks are invisible, then they are very typically removed by compression systems as a side-effect.

Obsolescence

When standards and formats change, it may be difficult to transfer DRM-restricted content to new media. Additionally, any system that requires contact with an authentication server is vulnerable to that server becoming unavailable, as happened in 2007 when videos purchased from MLB.com prior to 2006 became unplayable due to a change to the servers that validate the licences.

When Microsoft introduced their Zune media player in 2006, it did not support content that uses Microsoft's own PlaysForSure DRM scheme they had previously been selling. The EFF calls this "a raw deal".

In April 2008, Microsoft sent an email to former customers of the now-defunct MSN Music store: "As of August 31, 2008, we will no longer be able to support the retrieval of license keys for the songs you purchased from MSN Music or the authorization of additional computers. You will need to obtain a license key for each of your songs downloaded from MSN Music on any new computer, and you must do so before August 31, 2008. If you attempt to transfer your songs to additional computers after August 31, 2008, those songs will not successfully play."

Other – SMS Language & Abbreviations

What Is It?

SMS language (also known as chatspeak, txt, txtspk, texting language or txt talk) is the English language slang used in mobile phone SMS. It is an abbreviated form of English known as a rebus. It is similar to leet and AOL speak. With predictive text input increasingly being used, it is becoming less common.

It developed to accommodate the small number of characters allowed (early SMS permitted only 160 characters), and as a convenient language – the small keyboards on mobile phones. Without practice, sending SMS messages can be time consuming.

The objective of txt is to use the least the number of characters needed to put across a comprehensible message. Hence, punctuation and grammar are largely ignored.

Overview

Single letters can replace words. Examples:

- be becomes b
- see becomes c
- are becomes r
- you becomes u
- why becomes y

Single digits can replace words. Examples:

- ate becomes 8
- - becomes 4
- to or too becomes 2

A single letter or digit can replace a syllable. Examples:

- ate becomes 8, so:
  - activate becomes activ8
  - great becomes gr8
  - mate becomes m8
  - later becomes l8r
  - plates becomes pl8s
  - skater becomes sk8r
- - or Fore becomes 4, so:
  - before becomes (combining both of the above) b4
  - therefore becomes der4

There are miscellaneous adaptations of characters. Examples:

- ss becomes $
- oo becomes %
- '-orr-' becomes '-oz'

  - - example, Sorry becomes Soz, and Tomorrow becomes Tomoz (further abbreviated to 2moz.)

Combinations of the above can shorten a single or multiple words. Example:

- your and you are both become ur

Characters and punctuation are removed to shorten messages:

- Vowels are removed such that the sequence of consonants remain and the word is still recognisable.
  - - example, between becomes btwn.
- Whole words may be omitted, especially articles.
- Punctuation may be removed; only period and exclamation marks are generally used. The space and capital letter is often omitted after a period.

"/" signifies abbreviation, such as "w/" – "with" and "s/t" – "something".

Other transcriptions of slang or dialect terms can be used if shorter than the original words, as in cos (with fewer letters than because.)

Examples

Combining the above "techniques" can shorten whole sentences. Examples are as follows:

Are you going to the pub tonight?

- ru goin pub 2nyt

Hi mate. Are you okay? I am sorry that I forgot to call you last night. Why don't we go and see a film tomorrow? (120 characters)

- hi m8 u k?-sry i 4gt 2 cal u lst nyt-y dnt we go c film 2moz (60 characters)

General List Of SMS Abbreviations

SMS abbreviations are the recognised abbreviations of textspeak. They often contain phonetic symbols substituted – whole sounds (e.g., u – you or 8 – ate), and elide vowels from the represented word, unless the vowels function as initials. Marks of punctuation are omitted entirely, though "/' is used to signify there are two parts to the word (e.g., w/o means without). Note: some abbreviations listed here are not SMS- or internet-originated, having been used in other forms of communication before SMS and the internet were invented.

0-9

- 404 – I don't know (adopted from the http error code 404)
- 4u – for you

A

- a3 – anytime, anywhere, anyplace
- aap – always a pleasure
- aar – at any rate

- aas – alive and smiling
- add – address
- adn – any day now
- aeap – as early as possible
- afaik – as far as I know
- afk – away from keyboard
- aisb – as it should be
- aka – also known as
- aml – all my love
- aota – all of the above
- asap – as soon as possible
- a/s/l – age/sex/location
- asl – age/sex/location
- at – at your terminal
- atm – at the moment
- ayec – at your earliest convenience
- ayor – at your own risk

B

- b4 – before
- b4n – bye – now
- bak – back at keyboard
- bau – business as usual
- bbiaf – be back in a few
- bbiam – be back in a minute
- bbl – be back later
- bbs – be back soon
- bc – because
- bcnu – be seeing you
- bf or b/f – boyfriend or best friend
- bfn – bye – now
- bg – big grin
- blnt – better luck next time
- bm&y – between me and you
- bol – best of luck
- brb – be right back
- brt – be right there
- bta – but then again
- btdt – been there, done that
- btw – by the way
- bcoz – because

C

- cam – web camera
- cas – cracking a smile
- cmiiw – correct me if I'm wrong
- cmon – come on
- cnt – can't
- cob – close of business
- cos – because
- cr8 – create

- crb – come right back
- crbt – crying really big tears
- cu or cya – see you (later)
- cua – see you around
- cul – see you later
- cul8r – see you later
- cwyl – chat with you later
- cyo – see you online
- cys – see you soon
- cth – cleaning the house
- cuz – cousin

D

- dl or d/l – download
- degt – don't even go there
- diku – do I know you?
- dqmot- don't quote me on this
- dno – don't know
- dnt – don't
- dts – don't think so
- dv8 – deviate
- dw – don't worry

E

- ebkac – error between keyboard and chair
- eg – example or evil grin
- ema – e-mail address
- emfbi – excuse me – butting in
- eod – end of day
- eom – end of message
- ezy – easy

F

- f2f – face to face
- f2t – free to talk
- fbm – fine by me
- fc – fingers crossed
- ficcl – frankly I couldn't care less
- fitb – fill in the blank
- fnx – thanks
- fomcl – falling off my chair laughing
- frt – for real though
- ftw – for the win
- fwiw – for what it's worth
- fya – for your action
- fyeo – for your eyes only
- fyi – for your information
- f9 – fine

G

- g – grin
- g2cu – good to see you
- g2g or gtg – got to go
- g2p or gtp – got to pee
- g2r – got to run
- g9 – genius
- ga – go ahead
- gal – get a life
- gb – goodbye
- gbu – God bless you
- gd – good often repeated to be gdgd
- gdr or gd/r – grinning, ducking, and running
- gf or g/f – girlfriend
- gfi – go – it
- gg – gotta go or good game
- gg no re – good game, no rematch
- giar – give it a rest
- gigo – garbage in, garbage out
- gj – good job
- gl – good luck
- gl/hf – good luck, have fun
- gmta – great minds think alike
- gna – gonna or going to
- goi – get over it
- gol – giggling out loud
- gr8 – great
- gr&d – grinning, running and ducking
- gt – good try or get
- gtg – got to go

H

- h&k – hugs & kisses
- h2cus – hope to see you soon
- h8 – hate
- hagn – have a good night
- hago – have a good one
- hand – have a nice day
- hf – have fun
- hhis – head hanging in shame
- h/o – hold on
- hoas – hold on a second
- hru – how are you?
- hth – hope this helps
- hv – have

I

- iac – in any case
- ianal – I am not a lawyer
- ib – I'm back

- ic – I see
- icbw – it could be worse
- idawtc – I don`t agree with that comment
- idc – I don`t care
- idk – I don`t know
- idts – I don't think so
- idunno – I don't know
- ig2r – I got to run
- iirc – if i remember correctly
- ilbl8 – I'll be late
- ily or ilu – i love you
- im – instant message
- imao – in my arrogant opinion
- imho – in my humble opinion
- imnsho – in my not so humble opinion
- imo – in my opinion
- ims – I am sorry
- inal – I'm not a lawyer
- indtd – it's a nice day today
- iow – in other words
- irl – in real life
- irmc – I rest my case
- iuss – if you say so
- iykwim – if you know what I mean
- iyo – in your opinion
- iyss – if you say so

J

- jb – just because
- jic – just in case
- jk – just kidding
- j00r – your
- jac – just a sec
- jic – just in case
- jja – just joking around
- jk – just kidding
- jmo – just my opinion
- jp – just playing
- jpp – just playing player

K

- k – okay
- kk – okay, cool
- kl – cool
- kiss – keep it simple, stupid
- kit – keep in touch
- kotc – kiss on the cheek
- kotl – kiss on the lips
- knim – know what I mean?

L

- l33t – elite
- l8 – late
- l8r – later
- lbr – little boys room (bathroom – male)
- ld – later, dude or long distance
- lgr – little girls room (bathroom – female)
- lol – laugh out loud or lots of love
- lqtm – laughing quietly to myself
- lsl – laughing so loud
- ltm – laugh to myself
- ltns – long time no see
- lylas – love you like a sis
- lyk – like

M

- m8 – mate
- mfi – mad – it
- msg – message
- mtf – more to follow
- mtfbwu – may the force be with you
- musm – miss you so much
- myob – mind your own business

N

- n – and (though often completely omitted)
- n2u – new to you
- n1 – nice one
- nbd – no big deal
- nm – never mind
- ne – any
- ne1 – anyone
- nfm – none – me or not – me
- nimby – not in my back yard
- nlt – no later than
- nm – nothing much
- np – no problem
- no1 – no one
- noyb – none of your business
- np – no problem
- nrn – no response necessary or no reply necessary
- nt – nice try
- ntu – new to you
- nvm – never mind
- nw – no worries or no way
- nwo – no way out
- nvm – never mind

O

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io

- omg – oh my god
- oic – oh, I see
- omw – on my way
- oo – over and out
- ooh – out of here
- ootd – one of these days
- op – on phone
- otb – off to bed
- otl – out to lunch
- otoh – on the other hand
- ott – over the top
- ottomh – off the top of my head
- otw – off to work
- ova – over
- o – a hug as in xoxo

P

- pcm – please call me
- pdq – pretty darn quick
- plmk – please let me know
- pls or plz – please
- pm – private message
- pmfi – pardon me – interrupting
- pmfji – pardon me – jumping in
- poahf – put on a happy face
- ppl – people
- prolly – probably
- prt – party
- prw – people are watching or parents are watching
- ptmm – please tell me more
- pu – that stinks
- pxt – please explain that

Q

- q – question or queue
- qik – quick
- qt – cutie

R

- rofl – rolling on floor laughing
- rotfl – rolling on the floor laughing
- rotflol – rolling on the floor laughing out loud
- rotfluts – rolling on the floor laughing unable to speak
- rp – roleplay
- rl – real life
- rme – rolling my eyes
- rsn – real soon now
- ruok – are you okay?

S

- s icnr – sorry, I could not resist
- s ig2r – sorry, I got to run
- sal – such a laugh
- say – smiling at you
- sbtsbc – same bat-time, same bat-channel
- sc – stay cool
- sete – smiling ear-to-ear
- sis – snickering in silence
- sit – stay in touch
- slap – sounds like a plan
- slp – sleep
- slpn – sleeping
- smhid – scratching my head in disbelief
- snafu – situation normal all fouled up
- so – significant other
- sol – sooner or later
- somy – sick of me yet?
- sotmg – short of time, must go
- soz or sry – sorry
- spk – speak
- spst – same place, same time
- ss – so sorry
- ssdd – same stuff, different day
- ssinf – so stupid, it's not funny
- str8 – straight
- stq – search the Web
- suitm – see you in the morning
- sul – see you later
- sup – what's up?
- syl – see you later
- sxci – sexy

T

- t+ – think positive
- ta – thanks a lot
- tafn – that's all – now
- tam – tomorrow a.m.
- tb – text Back
- tbd – to be determined
- tbh – to be honest
- tc – take care
- tgif – thank God it's Friday
- thts – that is
- thx, thnx, thnq or ty – thanks or thank you
- tia – thanks in advance
- tiad – tomorrow is another day
- tlk2ul8r – talk to you later
- tmb – text me back
- tmi – too much information
- tmot – trust me on this

- tmr, tmz or tomoz – tomorrow
- tmwfi – take my word – it
- tnstaafl – there's no such thing as a free lunch
- toy – thinking of you
- tpm – tomorrow p.m.
- tptb – the powers that be
- tq – thank you
- tstb – the sooner, the better
- ttfn – ta ta – now
- ttly – totally
- ttml – talk to me later
- tttt – these things take time
- ttyl or ttul – talk to you later
- ttys – talk to you soon
- tu or ty – thank you
- txtm8 – textmate
- tym – time
- tyt – take your time
- tyvm – thank you very much

U

- ugtbk – you've got to be kidding
- uktr – you know that's right
- ul – upload
- ur – your or you're
- uv – unpleasant visual
- uw – you're welcome

V

- vf – very funny

W

- w/ or w – with
- w/e – weekend or whatever
- w/o – without
- wam – wait a minute
- wan2tlk – want to talk
- wayf – where are you from?
- wb – welcome back or write back
- wb2my – write back to my
- wg – well good
- whteva or woteva – what ever
- wiifm – what's in it – me?
- wk – work or week
- wkd – wicked or weekend
- wombat – waste of money, brains and time
- wrk – work
- wrud – what are you doing?
- wt or wut – what
- wtb – want to buy

- wtg – way to go
- wth – what the hell or what the heck
- wts – want to sell
- wubu2 – what you been up to?
- wuu2 – what are you up to?
- wu? – what's up?
- wubu2? – what have you been up to?
- wuciwug or wysiwyg – what you see is what you get
- wuf? – where are you from?
- wuwh – wish you were here
- wwyc – write when you can
- wylei – when you least expect it
- wat – what

X

- x or xxx – kiss
- xlnt – excellent

Y

- ya – your
- ybs – you'll be sorry
- ygbkm – you've got to be kidding me
- ykwycd – you know what you can do
- ymmv – your mileage may vary
- yr – your
- yw – you're welcome

Z

- zzz – bored or sleeping

Planet Limited (Company Nr 5952219)
Level 1, 120 Eleventh Avenue, Tauranga, New Zealand
www.*planetlimited*.io