NetComm™

Everything's Connected

# The NetComm Introduction to Broadband

# Table of Contents

# Preface – Who is this for?

This material is intended for those who need to know more ABOUT networking, broadband modems, wireless, VPNs and the like. It is intended for tech-literate, but non-specialist, readers, including general and business computer users, resellers, and anyone else who wants to know more detail about current broadband terminology.

To 'know about' is not the same as 'to know'. To really KNOW networking and TCP/IP takes several years of university-level education. But if you're a computer user nowadays, you need to know what the phone support people mean when they ask you what your IP address is. And to even KNOW what an IP address is, you have to know ABOUT networking.

If you want to set up a wireless link, or connect to a VPN, you might need to know about some these topics as well. Of course, if the Connection Wizard works as intended, or if you have an I.T. consultant on hand, you might not need to know any of it. But if you need to know more, out of necessity or curiosity, this will provide you with a comprehensive overview of the current technology involved in broadband networking.

Jonathan Shearman | Manager Affinity & Online | NetComm Limited | July 2004

# Chapter 1: Just Enough Networking

Once the preserve of specialists and technicians, computer networks are now very much part of everyday life, especially with the growth of broadband. And while you still might need to call in a specialist for initial setup or extensive re-configuration, it is increasingly necessary to understand the basics of network terminology and layout just to get by in the broadband environment.

Hence, ‘just enough networking’ This chapter will provide an introduction to the main concepts behind current Local-Area Networking (LAN) technologies such as:

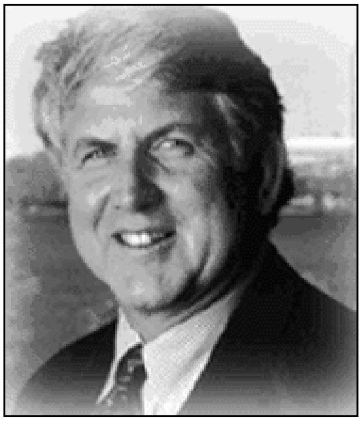
- Ethernet – its development and terminology
- Network topology
- Introduction to ‘protocols’
- Common networking terms and devices
- Introduction to the OSI model
- introduction to data packets and packet switching

Subsequent chapters will cover the basics of understanding file sizes and network speeds, wireless networking, hotspots, and VPN.

## Ethernet 101

Computers nowadays most often communicate by way of **Ethernet**[1]. In the last few decades, many kinds of networking technologies have come and (largely) gone; computers may have been connected to token-ring and AppleTalk networks by co-axial or AppleTalk cables. Nowadays, the industry standard is Ethernet on UTP cables, which is the focus of this article.

### History of Ethernet



Ethernet was pioneered by Robert Metcalfe (pictured) at the famous Palo Alto research centre of Xerox which gave birth to the laser printer and the graphic user interface that was to be adapted to the Apple Macintosh in 1984.

In 1973, Xerox was in the unique position of having to connect several hundred Alto computers. 'Unique' because at the time this was the only place in the world where several hundred computers existed. The aim was to provide a method to print to the newly-developed laser printer, which offered unheard-of speeds of one page per minute, and to do so from any one of the many computers in the building.

Originally dubbed 'the Alto Aloha Network' in honour of Norm Abrahamson's ground-breaking wireless work at the University of Hawaii, the term 'Ethernet' was eventually chosen as a reference of the early idea that the atmosphere is pervaded by a kind of 'æther' which provides the medium of transmission for light waves.

Xerox' initial development effort was later joined by Digital Equipment Corporation and Intel on Metcalfe's urging. (Subsequently, Metcalfe co-founded 3Com, now a $50 billion networking products company, and the first to market with an Ethernet card.)

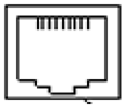
### Ethernet Becomes a Standard

Ethernet became a standard method of connecting computers in a local-area network and is defined by **IEEE standard 802.3**. This definition varied only slightly from the scheme perfected by Digital, Intel and Xerox, then known as '**DIX Ethernet**'. Subsequent refinements include **10Base2**, **10Base2**, and **100BaseT**, each defined by a distinct IEEE standard.

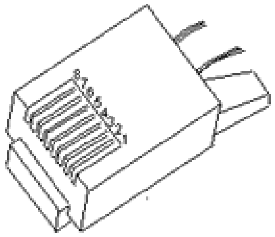
When people talk about LAN cables, Network cables, and Ethernet cables, these generally all refer to the same thing, all of them descendants of the original DIX Ethernet standard. Although Ethernet can be carried via Fibre-optic and co-axial cabling, here we'll concentrate on UTP (unshielded twisted pair) networking.

The figures below illustrates what current standard RJ45 Ethernet sockets and plugs look like, and how an Ethernet cable generally connects to a wall socket. You computer might already have a port like this, and if not, you'll need to buy an NIC – a Network Interface Card – which drops into your PCs PCI Slot, to provide a socket into which the Ethernet cable can be plugged.

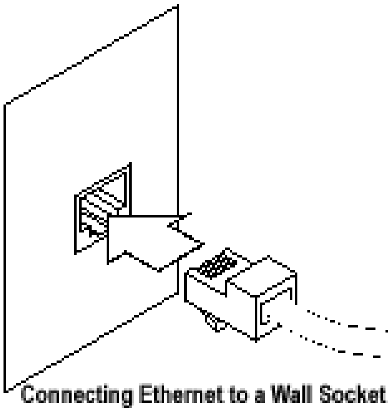
[1] See also http://en.wikipedia.org/wiki/Ethernet



RJ 45 Ethernet Socket



RJ 45 Ethernet Plug [not to scale]



Connecting Ethernet to a Wall Socket

## Cable and Connection Types

You will often hear the term '10BaseT' and '100BaseT'.

The initial numbers refer to the data transmission rate (=10 or 100 Mbps).

The term 'BASE' refers to '**baseband'** which describes the transmission of bits without further modulation, as distinct from '**broadband'** which transmits multiple signals through 'frequency division multiplexing'.

The letter 'T' stands for Twisted Pair (= the kind of cable used for this system).

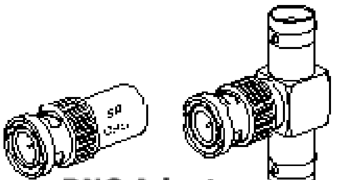
**10BaseT** carries 10 Megabits (not **megaBYTES**) per second and was described by the original definition of Ethernet. This is carried on twisted-pair cable similar to phone cable and classified as Category 3 cabling.

**100BaseT**, or Fast Ethernet, formalised in 1995, carries 100 Mbits per sec across Category 5 twisted pair cabling.

**Gigabit Ethernet**, finalised in 1998-99, handles 1000Mbits per Sec and can also be carried on Cat 5 cabling however Cat5e is preferred.

**Auto-Sensing:** Many Ethernet switches, hubs and routers are **auto-sensing**, automatically detecting and adjusting to the rate in use on the network to which they are attached.

Older forms of Ethernet were carried via co-axial cable, similar to that used for television aerial connections, and ending with a BNC connection. These were 10Base2 and 10Base5 formats and are rarely installed nowadays although there are still plenty in use. Also known as 'thinnet'.



BNC Adaptors

Ethernet can also be carried by fibre-optic cable which will not be discussed here.

**CSMA/CD**

As Ethernet is a 'shared bus topology', every device connects via the same cable. So each device 'listens' to the electrical activity on the cable and will only 'speak' - send a frame of data - when every other device is silent. This technique is called Carrier Sense on Multiple Access Networks.

But collisions can still occur when two devices commence transmission at the same time. When this occurs, both the devices will detect the collision and wait **for a random interval** before trying again. The **back-off algorithm** which handles this process is called the Collision Detection, or CD, response.

Together, these systems comprise CSMA/CD – Carrier Sense on Multiple Access Networks with Collision Detection – one of the century's great inventions, and one which lies behind nearly all modern PC networking.

Ethernet networks are often said to be divided into 'collision domains' which are network segments shared by a number of computers.

**Understanding Protocols**

In understanding networking, a term which comes up time and time again is '**protocol'**. A protocol simply establishes a means of communication. The simplest way to understand it is to compare it to the process of making a phone call.

First you need to know the number to call. Then you have to lift the handset and listen for a dial-tone. The phone you're calling has to be available and has to be answered. Then you need to establish your identity and that of the person to whom you are speaking and, if necessary, ask for the person you want to speak to. Only when all of these steps are completed can the telephone call proceed.

Network protocols are similar in principle: they are ways of establishing communications between different devices, different computers, different networks, and so on, following a set of steps understood by all parties to the conversation. So think of it as 'establishing a means of communication'.

## Local Area Network Topologies

**Types of Network**

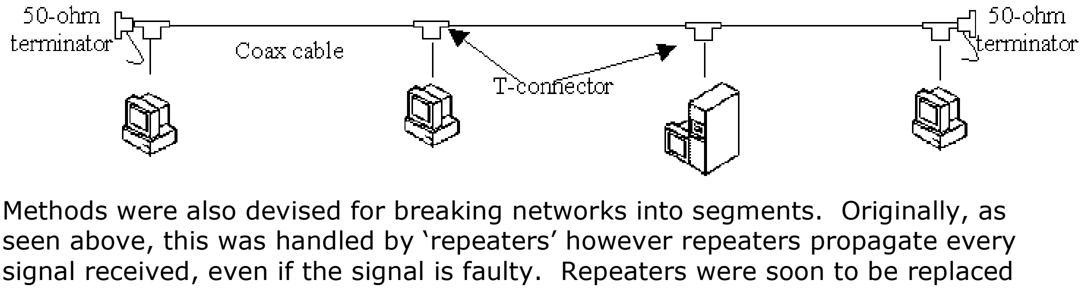
Networks are defined as Local Area Network (LAN); Wide Area Networks (WAN), which include the Internet and networks that conjoin offices in different geographical areas; Metropolitan Area Networks (MAN), and Wireless Local Area Networks (WLANS). The following is mainly concerned with LAN.

**Topology - Definition**

'Topology' refers to the kind of shape used to lay out the network and is used in two senses. A network's **physical topology** is the layout of the network cabling; the **logical topology** is the way signals are carried on the cable.

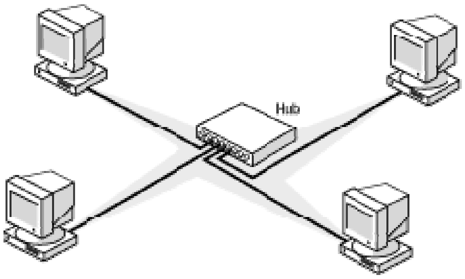
**Co-axial Bus Topology**

Original Ethernet networks were configured in **co-axial bus technology** – the computers were all attached to a straight length of cable. **Terminators** were installed at the end of cable lengths to prevent the signal bouncing endlessly up and down the wire. The limits were set by the distance the electrical signal could travel down the wire, which in this technology is 185 meters; different segments can be attached by **repeaters**, which pick up and amplify the signal.



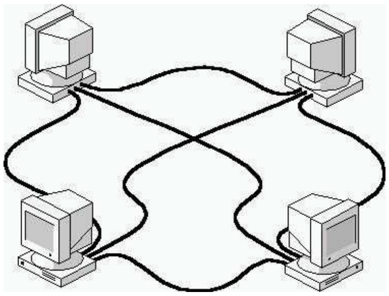
Methods were also devised for breaking networks into segments. Originally, as seen above, this was handled by 'repeaters' however repeaters propagate every signal received, even if the signal is faulty. Repeaters were soon to be replaced by **bridges** which are smart enough to only forward valid data frames, thereby enabling networking errors to be confined to the network segment from which they originated. (More on networking devices below.)

### Star Topology



In 10/100BaseT environments, the common network topology is star topology, with one or more workgroups connected to either hubs  or to routers which can in turn be interconnected (see below). Broadband home connections for more than one computer will often be based around a central modem/router which functions as a hub, and will resemble this kind of setup. (The difference between hubs, routers and switches will be discussed later.)

### Ring and Mesh Topologies



Ring topologies, as the name implies, link the connected devices in a loop of cable. This configuration was more commonly used in Token Ring networks, an IBM networking technology which is not discussed here.

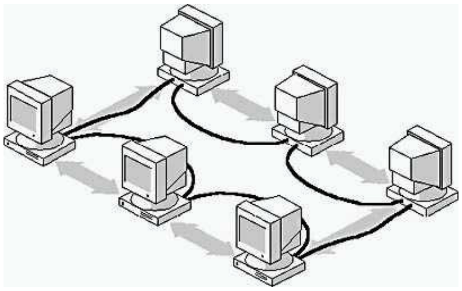
A mesh topology (right)  connects every computer to every other computer.  It is very secure from a fail-safe perspective however not often encountered as it generates a lot of connections and consumes a great deal of cable.

### Hybrid Topologies

Networks can be created using variations and hybrids of the common topologies above. These include **Star Bus** (no, not a coffee shop) where several network hubs are linked in a bus (straight-line) configuration with hubs branching off the support workgroups,  and **Star Ring** setups where the hubs in a star ring are connected in a star pattern by the main hub.

### Peer-to-peer Topology

Using peer-to-peer file sharing capabilities in Windows and Mac OS, it is easy to connect small networks of computers in a peer-to-peer network with no central hub as per the illustration below.  This can also be combined with Internet Connection Sharing to provide workgroup access to the Internet via a dial-in or broadband modem. Peer-to-peer is particularly common in home-user environments.



## Elements of Ethernet

### The Network Interface Card

The first link in the Ethernet chain is the Network Interface Card or NIC.  A basic 10/100 NIC such as the NetComm 1100 retails for around $25.00.  IT managers often choose a commercial-grade NIC such as the NetComm 1500 for installation into file servers; these retail for around the $100.00 mark and offer more network-management features and a more robust overall design.

Previously many NICs were hybrid models with various combinations of 10/100baseT, AUI and BNC connector, however with the move towards the 10/100 RJ45 format, this is what is most often seen nowadays.

### The MAC Address

Every NIC card has a unique serial number burned into it in electronic form. This is called the MAC (=Media Access Control) address (NOT to be confused with a Macintosh computer!)

The MAC address plays a pivotal role in directing network traffic as it is one of the fundamental values that enables data to be addressed from one machine to another. Later on we'll look at where, in the hierarchy that defines networks, the MAC address exists.

It is called 'Media Access Controller' because that is what it does. The network cable is the physical 'media' that carries the data to and from the computer. The NIC card controls

Next we'll look at the little black boxes that are found inside the wiring cabinets of computer networks. A word of caution – there is a degree of convergence going on between these devices, so that the difference between switches and hubs, and between bridges and routers, are becoming less and less clear over time.

### Segments

Trunks of cabling which connect devices back to the routers.

### Backbone

The backbone of a network is as the name implies the backbone cabling to which the main routers are attached.

### The Hub

As seen above, the first kind of device created to boost network performance was the **repeater** which amplified a signal weakened over long wires. Repeaters were characteristic of the co-axial-type 10Base2 and 10Base5 networks arranged in linear bus topology.

The **hub** was originally devised as part of the star topology characteristic of twisted-pair, 10BaseT networks. It incorporates the abilities of a repeater.

A hub distributes all of the data it receives through every port it has. Data in, data out. Basic hubs can only be used on one network segment; they have no ability to manage traffic between different network segments, or between different networks.

Hubs operate at the Physical Link layer which means they receive and transmit electrical impulses, with no ability to read the data in the packets which they transmit.

Other types of hub include **managed hubs**, which support SNMP [=Simple Network Management Protocol] and **stackable hubs,** which can be stacked to multiply the available network ports.

### The Switch

Switches are used to link several LANs, or to manage traffic in a busy LAN. Switches operate at the Data Link layer of the Network Stack, which means that they can read each data packet's MAC address.

A switch has multiple ports, each of which can support either a single workstation or an entire Ethernet LAN. A switch with a different LAN connected to each ports can switch packets between LANs as required.

Switches can be used in heavily-loaded networks to isolate data flow and improve performance.  For example, in a switch, data between two lightly used computers will be isolated from data intended for a heavily-used server. Alternatively, "auto sensing" switches that support both 10 and 100Mbps connections isolate 10Mbps traffic from the 100Mbps traffic.

Switches incorporate the capabilities of a transparent bridge to 'learn' the addresses of the computers to which they are connected by 'listening' to each computers' hardware [MAC] address. If a packet's destination is a station on the same segment where it originated, it is not forwarded. If it is destined for a computer on another LAN, it is connected and forwarded to a different port.

NetComm's Switch products include the NP2005 five-port, NP2008 eight-port, NP2240 24-port and the NP3005 five-port Gigabit Desktop Switch.

### The Bridge

A Bridge is a device that connects two networks so that they function as a single whole.  Bridges can also connect dissimilar network types, such as wireless and wired, or Ethernet and Token Ring.  Switches, mentioned above, incorporate many of the abilities of a bridge.

Note the Switches and Bridges both operate on the data link layer: this means they can read the MAC address of the packets they handle, but not the IP address, which belongs to a different 'layer' of the network stack, namely the Network layer, which can be read by routers.

### The Router

Routers are used to connect multiple LANs, or LANs and WANs: for example, to connect your LAN to the Internet.

The most sophisticated network connectivity devices, routers forward data packets according to their IP address, not according to the sending/receiving devices  MAC or hardware address.  Routers work by referencing a **routing table** which contains detailed information about the source and destination devices and the networks that need to be traversed.

IP Routers are **protocol-dependent** and are able to respond to variable factors in the environment to re-route around non-available network nodes, where required, according to TCP/IP rules.  The router helps determine the quickest route to the destination, taking into account changing network conditions such as network congestion or node failure.  This is managed by the RIP protocol[2] which is dynamically updated in response to constantly-changing network conditions.

Routers are nowadays small, portable devices that can be purchased for a few hundred dollars, or less; however, when they were first devised, they were large computers in their own right.  Cisco Systems started the commercial manufacture of routers in 1983 and it is only during the last several years that devices with routing ability have been available on the consumer market.

## The Seven Layer Model of OSI

During the development of networking from the 1960's onwards, a model was created to accommodate the widely varying kinds of systems which networked information services needed to traverse.  While we now take for granted that computers of all kinds are able to connect through the same kinds of cable and

[2] RIP (Routing Information Protocol) A simple routing protocol that is part of the TCP/IP protocol suite. It determines a route based on the smallest hop count (=least number of devices) between source and destination.

access common information sources, computer systems in the 1960s were far more diverse and nearly all proprietary, with completely different data and connection models in many cases. Gradually, with the emergence of standards and innovations such as TCP/IP, a model called OSI began to emerge which provided a vendor-independent model for networking.

OSI stands for Open Systems Interconnect. The purpose of the OSI model is to provide a descriptive framework for sending information across all of the different layers of a network – physical, logical, and electronic. OSI is a hierarchy that loosely defines how information is passed from the application layer, the topmost layer, through the other layers, down the network cable, across the internet, to the receiving computer and then back up to the receivers application.

Consider the composition and sending of an email. The familiar window into which you enter your text, the electronic address book from which you choose your recipients, and the controls that allow you to Send and Receive are all features of the particular application you use for email. Yet when you're done writing and you want to send, you click on a button, and away goes your info, through a set of cables which have none of features you have just used, in the form of a swarm of data packets, which are ultimately just ones and zeros, which are ultimately just electrical pulses, or flashes of light, or sound waves, on a length of cable.

Between the sender and the receiver, the data is converted into packets; it might be encrypted for security; then it is sent through one, or a number, of routers; then sits on a computer waiting for the receiver to connect and download it, before turning up as an Unopened Item in the receiver's Inbox with all of the formatting that you gave it when you sent it. While in transit, different protocols are invoked at various points in the process, including SMTP, TCP/IP, and many others that a computer user doesn't need to know about.

The table provides a summary overview of the 7 Layers of OSI.

| LAYER | FUNCTION | PROTOCOLS | HARDWARE |
|---|---|---|---|
| Application Layer<br>IS LIKE:<br>Car dashboard providing user-level information and feedback | **Mediates application processes and network services.** Application software specific. Manages communication between applications, data flow and error recovery. | FTP (File Transfer)<br>SNMP (Simple Network Management)<br>SMTP (Simple Mail Transfer), NCP (Network Control) | Computer/Gateway |
| Presentation Layer<br>IS LIKE: traffic direction signs on highway | **Translates to/from application to network format** – handles protocol conversion, data translation and encryption, and data compression. Special application called Redirector operates here | NCP | Computer/Gateway |
| Session Layer<br>IS LIKE: Traffic signals or traffic police | Provides synchronization between computers; allows applications on two computers to establish a session. | - | Computer/Gateway |
| Transport Layer<br>IS LIKE: speed cameras and gauge meters | **Ensures data transmission accuracy**. Ensures reliable delivery of packets. Repackages messages, dividing them into smaller packets. | TCP (Transmission Control)<br>SPX (Sequenced Packet Exchange)<br>NWLink (Microsoft's implementation of IPX/SPX)<br>NetBEUI | Computer/Gateway |
| Network Layer<br>IS LIKE: Exit Ramp Numbers | **Establishes THE UNIQUE NETWORK ADDRESS AND MANAGES THE TRANSPORT OF INFORMATION PACKETS BETWEEN DIFFERENT NETWORKS.** Routes the packets and determines the best route for sending them. Manages network traffic | IP (Internet Protocol)<br>IPX (Internet Packet Exchange), from Novell Netware<br>NetBEUI (NetBios Extended User Interface) legacy code from MS-DOS<br>DLC (Data Link Control) | Routers |

| | | | |
|---|---|---|---|
| | problems, packet switching, routing, and reassembling data. | | |
| Data Link Layer IS LIKE entry ramp/toll booth that allows access onto the highway | **DETERMINES HOW DEVICES CONNECTED WILL GAIN ACCESS**. Translates binary values into frames/packets. Sends data frames from the Network Layer to the Physical Layer. Contains the physical address of a device. **Like the FROM: and TO: of a envelope**. | None | Bridge, Remote Bridge Switches<br>IEEE 802.X Standards<br>Ethernet, Token Ring, FDDI |
| Physical Layer IS LIKE the actual material from which road is built. | Transmits data bits between computers. Defines how 1's and 0's are interpreted from the medium. **DEFINES CABLES, NETWORK INTERFACE CARDS, AND OTHER PHYSICAL ASPECTS**. | None | The NIC CARDS, REPEATERS, CONCENTRATORS/HUBS, SWITCHING HUBS etc.<br>CABLES - IEEE 802.X, etc |

## IEEE 802.X SPECIFICATIONS

THE INSTITUTE OF ELECTRICAL AND ELECTRONIC ENGINEERING (IEEE) standards relevant to ethernet and wireless networking are:

- 802.1 Internetworking
- 802.2 Logical Link Control (LLC)
- 802.3 CSMA/CD LANs (Ethernet)
- 802.4 Token Bus LAN
- 802.5 Token Ring LAN
- 802.6 MAN – Metropolitan Area Networks
- 802.7 Broadband Technical Advisory Groups
- 802.8 Fiber Optic Technical Advisory Group
- 802.9 Integrated Voice and Data Networks
- 802.10 Network Security
- 802.11 Wireless Network

## PC Networking and TCP/IP

### The Relationship of Ethernet, TCP/IP and other network protocols

As computer technology developed in the 1960’s and 70’s, a number of companies were devising ways of networking their computers. Meanwhile, the Internet – then called ARPAnet – was being developed as a way of networking ANY computers. So these two efforts were, to some extent, at cross-purposes. Many companies were concerned with gaining a marketplace advantage by creating new capabilities; others, such as IBM, were sufficiently established to define the game on their own terms and intended to capture a significant commercial advantage by so doing.

As a result, a number of networking methods were developed, such as IBM’s Token Ring, Novell’s IPX/SPX, AppleTalk, Microsoft NetBEUI, and DECnet to name a few. None of these systems were designed specifically for inter-operation with other systems and all had varying degrees of longevity, the most successful being arguably Novell’s IPX/SPX, which remains in common use today. Some of the others are still in use but are not subject of many new installations.

As the industry evolved in the 80’s and 90’s, many computer users began to object to proprietary systems of all kinds on the grounds that they concentrated market power in the hands of the vendor and discouraged competition on a ‘level playing field’. ‘Open systems’ and ‘vendor-neutrality’ became catch-cries for those who were wary of being locked in to a proprietary computing environment and forced to purchase equipment from a particular vendor. Vendors also began to realise that an ‘open-systems’ approach might have its own commercial rationale. These were among the factors that drove an industry-wide migration towards publicly-defined standards such as Ethernet and TCP/IP as the basis of local-area networking.

The current state of play is such that nearly all new LAN installations are built around Ethernet standards, and most network operations are based on Microsoft, Apple or Unix implementations of TCP/IP.

### Packet Switched Networking: a Cold War Child

Why was Internet routing invented? Understanding this is key to understanding the basic facts about TCP/IP, the Internet, networking and the 'big picture'.

The Internet had its origins in a project called ARPANET which was started in 1968. It evolved to become the Internet during the subsequent decades. At the outset, the main sponsor was the US Department of Defence.

ARPANET was a child of the Cold War. The problem it was trying to solve was that of the effect of a nuclear attack on the USA. Any communications network would obviously be a priority target in such an attack, and so if there was a ‘central switchboard' facility it would only have to be targeted to take out communications all over the USA.

The Defence planners were acutely conscious of the strategic risk presented by a central communications facility. So 'Routing' and ‘Packet Switching' were devised to meet this challenge.

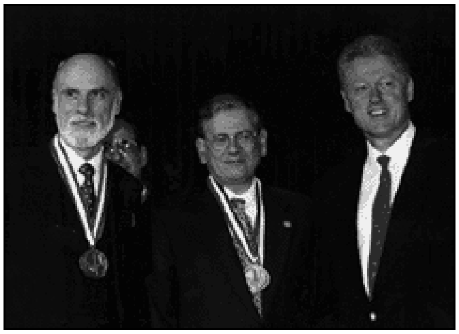
In a packet-switched network, messages are divided into ‘packets' and individually sent with their destination attached - rather like the address on an envelope. Each data packet contains part of the complete message and the 'routing' information that will enable it to reach its intended destination. As well as this, it contains the code which informs the receiver which message it is part of, along with code which enables the receiver to know if it has all been transmitted successfully. In the event that either (1) part of the route to the destination becomes blocked, or (2) part of the message is lost, the system will either re-route the data or re-send the missing components, or both.

When the data packets arrive at their destination, they are re-assembled to deliver the original message content. And this all happens in an instant.

This complex sequence of events is controlled by TCP/IP which stands for Transmission Control Protocol/Internet Protocol. This is the underlying technology behind the Internet and most current PC networking. You could say that it is to the information age what electricity was to the industrial revolution. It is an amazing technical achievement and one of the great inventions of the 20th century.

An important point is that TCP/IP was not invented for consumer convenience. It was designed as a military technology, by scientists with advanced degrees in mathematics and computer science. Now you're using it to go shopping and look up the train timetable. Hey, that's life, nowadays. So while some might complain that using the Internet ought to be as easy as making toast, this article assumes that it never will be, and encourages you to learn more about how it works. And no, this doesn't mean getting a degree in computer science, because you will probably never have to understand it that deeply. But getting an idea of how TCP/IP works is essential.

Below, two of the inventors of TCP/IP, Vince Cerf (left) and Bob Kahn, with President Bill Clinton, at an award ceremony recognising their contribution to the creation of the Internet.



## *Internet Basics 1: IP Addresses*

### What does TCP/IP stand for?

TCP/IP stands for Transmission Control Protocol/Internet Protocol. TCP/IP is a suite of software protocols and services that control the way information flows around the Internet. It is embedded in every current-model computer system and also in the routers that direct Internet traffic around the planet. Later we'll look at where in your computer you will find TCP/IP settings.

### What's an IP Address?

For any information to reach you via network, your computer must have an address. The address enables the network to distinguish and identify every computer that is connected to it, in much the same way that a postal address distinguishes every house and enables deliveries to be made.

Although it serves exactly the same purpose as a postal address, a computer's IP address is a number, which is formatted as a series of digits divided by points, for example:

192.168.1.1.

This is called an IP Address, meaning 'Internet Protocol Address'. Every device attached to the Internet has an IP address – not only PCs, but also PDAs, and Internet-enabled cameras and printers. This means that the pool of IP addresses must be a very large number, as there are millions of such devices with many more continually being added. Later on we'll look at issues involved with maintaining the pool of available IP numbers.

### The Anatomy of an IP address

Part of the IP address signifies the network identity – that is, the network which the computer is part of. Another part of the address signifies the host identity, where 'host' means an individual device attached to the network. These two components can be thought of as a street name and a house number.

An associated group of digits comprises the 'subnet mask'. This information is used to subdivide IP address information between varying numbers of devices - think of it as creating room numbers for each device. The Subnet Mask number usually appears as 255.255.255.0.

The combination of the IP address and subnet mask must always result in a unique ID for each device, otherwise information cannot be delivered accurately, meaning either that the network will fail, or that connection to the Internet cannot be made or will perform erratically, as network devices will not be able to accurately distinguish between devices that have the same address data.

Now if your computers were on their own network, and you only had a small number of them, then you would only need a very small range of IP addresses to distinguish one computer from the other; you could simply number them from one to six. And early network technologies were mainly designed on this basis, as, at the time, few foresaw a day when millions of computers would be interconnected around the world. Over the years, integrating these earlier network techniques with the requirements of IP took some very nifty re-engineering. However by 2004, virtually every desktop computer incorporates TCP/IP in the operating system. Before we look at the TCP part of the equation, we'll finish more detail on IP addressing.

### IP Numbering Rules

As we have seen, IP addresses are the numbers which identify every computer (and other kinds of device) on the Internet. But what are the rules governing IP addresses? Have you ever wondered why an ADSL modem's 'default' IP address is nearly always 192.168.1.1?

(Incidentally, all of what follows deals with IP version 4, which are current in most systems; Version 6, which is 'in the works', is not dealt with here.)

### IP Address Number Format

As we have seen, an IP address consists of a string of four 8-bit numbers, usually notated in decimal format, divided into four 'octets', for example:

192.168.100.123

which in binary format is 11000000.10101000.1100100.1111011 (click here for a refresher on Binary). The highest number of any of these groups of octets is 255, as higher values require another 'binary digit' or 'bit' to describe.

### What an IP Address Does

As we have seen, an IP Address is a number that uniquely identifies every IP network on the Internet, and every IP device on the network.

Part of the IP address identifies the network, and the other part identifies the individual devices ('hosts') which comprise the network.

As the IP number itself has to be divided between Network ID and Host ID, some consideration has to be given as to how the number is subdivided.

### IP Address Classes

If two octets of the IP address were allocated to the Network ID, and two octets to the Host ID, it would mean that the total number of IP networks could be 65,536, each with 65,536 hosts. Why? Because the each half of the IP address, comprising two octets of four numbers, spans this number of possible values.

The problem with this scheme would be (a) not enough possible network identifiers, and (b) a lot of address numbers would be wasted as most networks would have much smaller numbers of devices.

After a bit of thought, the inventors of IP addressing decided that there could be several ways of dividing up the digits, according to the size of the networks that would be needed. So the idea of IP address classes was devised, and the possible types of IP address divided into three - Class A, Class B and Class C (with two other types, D-E, reserved for special purposes).

In this scheme, the first four bits of the address detail are used to flag the type of IP address:

| Address Class | Leftmost bits | Starting at | Finishing at |
|---|---|---|---|
| A | 0xxx | 0.0.0.0 | 127.255.255.255 |
| B | 10xx | 128.0.0.0 | 191.255.255.255 |
| C | 110x | 192.0.0.0 | 223.255.255.255 |
| D | 1110 | 224.0.0.0 | 239.255.255.255 |
| E | 1111 | 240.0.0.0 | 255.255.255.255 |

### The Class A Address

The Class A address is designated for large networks. The first octet is the Network ID, meaning that only 126 such network IDs are available. However each Class A Network can have more than 16 million hosts. Only about 40 Class A Networks are actually in use, in companies and institutions including IBM, the Army Information Centre, Hewlett-Packard, Apple Computer and the Ford Motor Company. Example Class A address:

| Network ID | Host or Network Node ID |
|---|---|
| 114 | 24.54.103 |

### The Class B Address

The Class B is used for medium-sized networks for example University campuses. IP addresses with a first octet ending in values from 128 to 191 are class B addresses. The second octet is also part of the Network ID, while the remaining octets are used to identify each host. This means that there are 16,384 Class B networks possible, each with up to 65,534 hosts for a total of 1,073,741,824 unique IP addresses. Class B networks make up a quarter of the total available IP addresses. Example Class B Address:

| Network ID | Host or Network Node ID |
|---|---|
| 145.214 | 513.106 |

### The Class C Address

Commonly used for small business networks. First octet is generally in the range 192-223 with the next two octets also used for Network ID. This gives a total of more than 2 million network address ranges, with the remaining octet flagging Host or Node ID. In this scheme seems to provide only 255 IDs for hosts or nodes however this can be increased by Subnet Masking (see next).

Example Class C Address:

| Network ID | Host or Network Node ID |
|---|---|
| 197.224.132 | 106 |

Class D and E addresses will not be discussed here.

### Private Network Numbers

Within IP numbering, certain ranges have been declared for 'private' or intranet addresses. These ranges are as follows:

| Class | Starting at | Finishing at |
|---|---|---|
| A | 10.0.0.0 | 10.255.255.255 |
| B | 172.16.0.0 | 172.31.255.255 |
| C | 192.168.0.0 | 192.168.255.255 |

This is why many ADSL modems have a default address of 192.168.0.0. In fact, any network node or host is able to use a 'private' IP number provided they are not directly connected to the Internet - for example, if they are behind a firewall, or connected through a NAT device such as an ADSL router. If a network is connected through a router, the router will generally have both an 'internal' IP address and a 'Public' or 'WAN' IP address which has been allocated by your ISP.

When your NetComm router connects to the Internet via your ISP, it generally receives an additional IP address, which is the WAN or Public IP address that routes data between your computer(s) and the rest of the Internet.

This is the difference between 'Private' and 'Public' IP addresses, which is a really important distinction to understand for any networks connected to the Internet via broadband.

.

### How IP Numbers are Allocated: The IANA

Given that IP addresses must be different for each device, the question arises as to how this system can be maintained when there are so many millions of devices connected to the Internet, with more being added all the time.

The IP address system is administered globally by the Internet Assigned Numbers Authority (www.iana.org) which is 'Dedicated to preserving the central coordinating functions of the global Internet for the public good.' Operating under the umbrella of the Internet Corporation for Assigned Names and Numbers [ICANN], IANA delegates IP address registration services through a network of Regional or National Internet Registries [RIRs and NIRs].

IP address allocation for Australia is handled by APNIC, the Asia Pacific Network Information Centre, which also handles countries as diverse as Afghanistan and Korea. One of four Regional Internet Registries, APNIC is situated in Brisbane.

RIRs in turn delegate a great deal of day-to-day IP address allocation to Internet Service Providers or Local Internet Registries, which provide IP addresses on the 'retail' level for most small businesses and individuals. To qualify for this service, ISPs must apply for and receive membership in APNIC, thereby becoming part of the network which distributes IP address data worldwide. Of the blocks of IP addresses that are provided to ISPs, many are allocated dynamically through DHCP on a per-session basis as their end-users require, while others are reserved for clients to provide a 'static' IP address that does not change from session to session (for which an additional charge is usually made.)

### How Big is the Pool of IP Numbers?

Incidentally, even though the number of potential IP addresses is large - somewhere north of four billion - it is nevertheless finite, and very early in the days of the Internet, it was realised that an alternative method would eventually be needed to provide a larger pool of potential addresses, especially as more and more devices began to carry IP addresses (for example phones, PDA's, digital cameras, printers, and so on.)

The current IP address scheme is called IPv4. It will eventually be replaced by IPv6, which allows for a truly astronomical number of IP addresses - enough for every grain of sand on Earth, it is said. Widespread adoption of IPv6 is still some way off, but despite some alarmist articles, the real experts seem to be comfortable that there are plenty of numbers for the foreseeable future.

### Network Address Translation

As mentioned elsewhere, NAT is an IETF standard that allows an organisation to present itself to the Internet with far fewer IP addresses than there are machines on its network. The NAT technology, which is usually implemented in cable and xDSL routers, converts the private IP address of a machine on the internal private network to one [or more] public IP addresses for the Internet. It changes the packet headers to the new address and keeps track of each session. When packets come back from the Internet, it performs the reverse conversion to the IP address of the client machine.

This is why your personal computer(s) may continue with a 'private' IP address, such as 168.192.1.4, an IP address which is in use by millions of devices around the world: the NAT device accommodates the 'Public' IP address and translates address information between your PC and the Internet. Again, this is the difference, mentioned elsewhere, between Public and Private IP addresses.

## *Understanding TCP*

Getting a handle on TCP – Transmission Control Protocol – is another one of those things that is practically essential in the Computer Age. Again, there are a lot of levels of understanding of TCP, and a complete description occupies several large texts. The purpose of this article is to describe it in basic terms – just enough knowledge to form an idea of how it works.

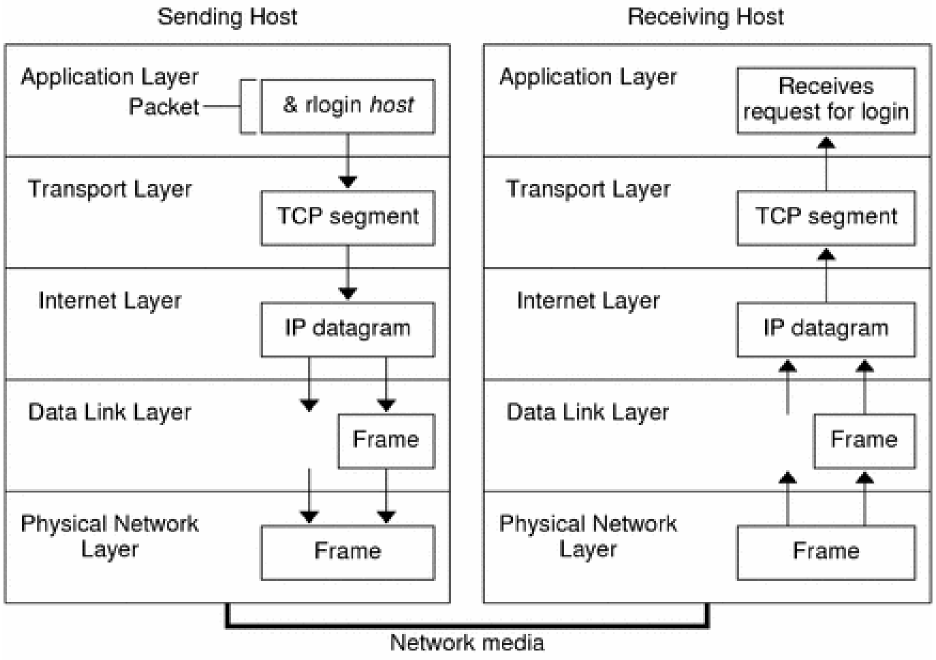
### What is TCP?

TCP is part of whole suite of standardised communications protocols that enable diverse kinds of computer systems to communicate across many kinds of media – from ethernet to fibre-optic to satellite to wireless transmission. Even though this suite includes many components other than TCP, the whole suite is known by the title 'TCP/IP'..

The meaning of the word ‘protocol' is very important here. Simply put, a protocol is a standardised way of establishing communications via a set of agreed procedures and rules. In context of TCP, the protocols define the set of standards that a computer's hardware and software must follow in order to be recognized and understood by other computers. Computer hardware and software makers incorporate the protocol into their devices, for if they are ‘TCP/IP – compliant' then they will be able to interact in predictable ways with other devices.

And because this protocol has to work through a number of ‘layers', from the ‘topmost' layer where the user enters information into an application displayed on a monitor, right down to the bottom layer, where the information has been transformed into a series of electrical, optical or radio pulses for transmission, the entire suite of protocols is often referred to as a ‘stack'. Hence, the ‘TCP/IP

Protocol Stack', as shown hereunder.



Let's look at the layers of this stack.

### The Application Layer

The topmost layer, that part of the TCP stack nearest the user, is the Application Layer, and is where user software resides. This is where applications for file management and Internet applications such as Outlook and Explorer operate.

Programmers need only concern themselves with the standard methods that have been defined to allow software on the Application Layer to interact with processes on the lower layers through an API, or Application Programming Interface. The API provides a standardised way of dealing with the lower levels of the Network Stack. Examples of application-layer protocols include:

**HTTP –** the protocol that serves Web pages

**SMTP –** email protocol

**MIME** - Multipurpose Internet Mail Extensions; more of a standard than a protocol. Used by e-mail applications, it is a method for sending attachments with electronic mail.

**PING**—A command used to verify the existence of and connection to remote hosts over a network

**TRACERT**—A diagnostic utility that displays the route a packet has taken to a destination

**FTP**—A protocol for transferring files to and from a local hard drive to an FTP server located on another TCP/IP-based network

### The Transport Layer

Data processed by an application on the top level is then passed via the process defined through the APIs to the Transport Layer where it is prepared for transmission across the Network.

The Transport Layer is a major part of the unique communication method of the Internet. To traverse the Transport Layer, the information being sent is sliced into individual pieces, called Packets, and each is stamped with the IP Address of the Sending and Receiving device.

Packet-switching is at the heart of Internet technology. As discussed in The Internet as a Cold War Child, the idea of slicing information into ‘packets’ each of which can travel independently, is **the** central idea behind the Internet.

A data packet is a unit of information, with the text and graphical elements converted into binary format and enclosed in a sequence of other values which control the sending and receiving process. The data content of the packet is referred to as the ‘payload’.

TCP Data Packets always have the same format, with the ‘Sending’ and ‘Receiving’ information represented in a consistent way and at the same point in the sequence. Other parts of the packet indicate which message the packet is part of, when it was sent, and special code to enable the receiver to confirm that the data has been received correctly. These processes are referred to as flow control, verification and error correction.

TCP is referred to as a ‘connection-oriented protocol’ which establishes and maintains a connection between remote machines for the duration of a transmission.

The Transport Layer also has another protocol called UDP, for User Datagram Protocol, without error-correction and verification. A so-called ‘connectionless protocol’, UDP is often used in streaming media, conferencing and VoIP applications.

In summary, transport Layer protocols in include TCP and UDP.

### Internet layer

Beneath the Transport layer is the Internet layer. Three key protocols reside in the Internet layer: Internet Protocol (IP – as discussed previously), Address Resolution Protocol (ARP), and Internet Control Message Protocol (ICMP), along with two less-used protocols, Reverse Address Resolution Protocol (RARP) and Internet Group Management Protocol (IGMP).

Address Resolution Protocol (ARP) matches IP address requests, such as www.netcomm.com.au, with a DNS number, such as 216.218.210.207.

ICMP is part of the TCP stack that handles verification, correction and resend requests.

RARP and IGMP won't be detailed here.

### The Data Link Layer

The next layer is the Data Link layer. Having been 'packetised' and 'addressed', the data is then 'encapsulated' in a format suitable for transmission across the network. 'Encapsulation' re-formats the data again into a format compatible with lower-level network protocols, and adds another layer of information to identify and demarcate data packets as they traverse the network. 'Encapsulation' places 'packets' within another type of packet, often called a data 'frame'.

Examples of a data-link-layer protocols include Ethernet, FDDI, Token Ring, and Wi-Fi.

**The Physical Layer**

The bottom layer of the stack is the layer on which network cards and card driver software operates. Among other things, Physical Layer services interface with the computer network hardware and provide another level of error correction code to ensure that data is transmitted correctly on the physical-link level.

**Summary**

This brief description has introduced

- the concept of the TCP/IP Stack

- Some essential attributes and applications of each network layer

- Some key protocols belonging to each network layer

This article has introduced IP addressing and the main components of TCP. Of course, these are big topics, but the emphasis here is on getting just enough knowledge to start forming a meaningful idea of how they work. Once the idea of data packets being routed across networks by their IP addresses becomes familiar, a lot of things about the Internet become a lot easier to grasp.

## *Understanding DNS*

If you understand the basics of IP addresses and TCP/IP, the remaining idea required to get a grasp on the basics of the Internet is the DNS. DNS is an abbreviation for Domain Name System, but it can also mean Domain Name Server, depending on the context in which it is used.

DNS is an heirarchical naming system that allows computers to be defined and located through user-friendly names rather than through hard-to-remember IP numbers. The DNS system does this through resolving DNS addresses to their IP number equivalent. It was devised by Paul Mockapetris, then a member of the IETF.

When a user enters a website address in their browser, the DNS sytem rapidly matches the registered domain name, say www.netcomm.com.au, with a matching IP address, say 216.218.210.207, through a procedure called DNS lookup. The IP address it looks up will usually be that of a computer server which responds by sending Web page information through the HTTP protocol. Such servers are generally called webservers.

What actually happens when the user enters the address is that the user's computer 'asks' the nearest name-server 'What IP address corresponds with this Domain Name?' If the computer that is asked does not know, it passes the query up to the next name server, until it reaches an 'authoritative source' which is a computer that knows without having to ask. The ultimate authorities are the Root Servers, mentioned below.

Although this is generally quick, it can sometimes take 15-20 seconds to resolve an address request, during which time the browser will display an 'attempting to find address' message; any longer than this and the web browser is likely to give up and declare the address unreachable.

### How DNS is co-ordinated and maintained

Maintenance and administration of the DNS falls under the responsibility of ICANN, the Internet Corporation for Assigned Names and Numbers. ICANN in turn devolves responsibilities for the sale and registration of Domain Names to accredited registrars, a list of which is here. If one wishes to register a domain name, these are the companies to which application must be made.

For the DNS system to work, a current list of all registered domain names must be accessible to any computer connected to the Internet. This bit of Internet wizardry is accomplished by distributing copies of DNS records to 13 'root-servers' located at various points around the planet. The data contained in these records is then propagated throughout the Web and is 'cached', or kept in the short-term memory, of computers at Internet Service Providers worldwide. A variety of other techniques are used to ensure that DNS lookups happen very quickly from your average desktop.

Incidentally, every shipping PC has the root-server information built into it at the factory. This is one of the reasons why the Internet works so seamlessly.

As an illustration of the way the DNS works, note that when you register a domain name or move it to a different hosting service, it takes 24-72 hours for the change to ripple through the network of Root Servers and thence through the remaining records on the Internet.

### Components of a Domain Name

We have seen that when we enter a web address, the DNS matches the request with the IP address of a Webserver; if it can't make the match, it will return a 'can't find server' error message.

Let's look at elements of a web address in more detail.

### Top Level Domains

The web address frequently ends with a gLTD, or Generic Top Level Domain. This is an extension, such as .com, .org, .net. and so on, which loosely describes the kind of organisation that the website belongs to. These are summarised as follows:

| | |
|---|---|
| .aero | air transport industry |
| .biz | business use |
| .com | general commecial |
| .coop | co-operative associations |
| .info | Informational resource sites |
| .museum | museums |
| .name | individuals |
| .net | various networks including some some broadcasters [e.g.www. abc.net.au] |

| | |
|---|---|
| .org | intended for non-commercial organisations |
| .pro | credentialed professionals and professional bodies |
| .gov | Government exclusive |
| .edu | accredited educational institutions |
| .mil | US Military |
| .int | registered organisations established by treaty |

Some of these extensions are rarely encountered; almost every site encountered is likely to be .com, .net., .org, and .gov with .biz sites seen from time to time. It is interesting to note that, for some reason, in the UK, .com sites generally have the gLTD of '.co'.

**Country Codes**

Many web addresses have an additional TLD code which signifies their geographical location, for example, .au for Australia, .uk for Britian, .ca for Canada, and so on; a complete list is here. Note that the gLTD '.com' can be registered anywhere in the world; it is a common misconception that '.com' sites must be registered in the US. In fact in the early days of the Web, any site with no country indicator could be assumed to be a US site, but since 2002 the TLD '.us' has existed to denote US websites, although it is not often seen.

**Second-Level Domain Name**

The second-level domain name comprises the unique identifier that describes the company, organisation or network that owns the domain, for example 'NetComm'.

**Subdomains**

You might have different subdomains, indicated by the left-most sequence of letters in a Web address. The default value is 'www' however this is often changed to denote a particular function or division with a domain, for example <support.companyname.com> for company support information, or the site hosting this document which is <lc.netcomm.com.au>. Subdomains also allow companies to run diverse kinds of webservers utilising different server technologies under the umbrella of one domain name, among other things.

**Hostnames**

Finally a web address might include a hostname which is the name of a specific computer on a network. A Fully Qualified Domain Name will include full descriptions of all of these elements of a Domain Name.

Next we will look at some other important aspects of understanding DNS.

**Finding Out Who Owns a Domain Name**

A simple search of the ownership details of any of the currently-registered 43-million-odd domain names in the world can be conducted through a 'whois' search via any of the ICAAN registrars for the area in which the website is owned, for example at whois.melbourneIT.com. for Australian records.

When such a search returns a result, you are seeing details of the A-Record or Address Record for a DNS entry. This will contain contact information for the domain owner, in addition to the all-important Name Server record which directs browsers to the appropriate Domain Name Server when a Web Address is entered.

**DNS Tricks and Techniques**

The structure of DNS allows for a lot of flexibility in the way websites are hosted. For example, a virtual web server can accomodate any number of Domain Names on a computer with one IP address. Many hosting services exploit this principle to host multiple sites on their machines.

Conversely, one domain name may map against a number of IP addresses which is useful for fail-safe and load-distribution purposes.

DNS also includes:

- MX Records, for Mail Exchange, used for specifying how domain email should be routed

- CNAME record or canonical name record makes one domain name an alias of another. The aliased domain gets all the subdomains and DNS records of the original

- NS record or name server record maps a domain name to a list of DNS servers for that domain

- SOA or Start of Authority record for specifying the DNS server and providing authoritative information about an Internet domain

(This is not an exhaustive list.)

Zone Record

The extent of the Domain is indicated by a Zone Record, which may sub-divide the Domain into several zones or against a number of IP addresses.

**DNS and the URL**

We have already seen how a web address is resolved against an IP address by DNS. However in many cases a Web address will contain an additional layer of detail to the right of the basic Web address, separated by a '/'.

This indicates a URL, or Universal Resource Locater. Essentially a URL is a request for a particular file, directory, or process, on a named server.The URL is simply the web address in question, along with an indicator of the particular file on the server.

At its simplest, a URL simply fetches a file in a named location, for example, the URL that fetches the file you are reading is <lc.netcomm.com.au/DNS_2.htm>

If you enter a web address/directory name in your web browser, the browser will generally attempt to find and display a file called 'index.html' or 'default.html' located within the directory. If it can't find one, it will display an error message.

An URL can also be 'active' for example when it includes a procedure to execute an Active Server Page [.asp] or .php script. In these cases, the webserver does more than serve a web-page - it executes a 'server-side' process which may interogate a database and return a record, update a user registration, or any number of other processes.

# Chapter 4: How Fast is 'Fast?'

Computer users will frequently read about how 'fast' broadband modems are, compared to 'old-fashioned' dial-up modems. But this leads to the question – how fast does my connection need to be? Just how 'fast' is 'fast', and how fast a connection do I need?

To get some perspective on this, realise that modern data transmission technologies are very, very fast indeed. There are technologies that can transmit gigabytes of data across a continent in a matter of minutes. Obviously, no individual is going to need this performance. At the other end of the scale, you will often want to download music files or software updates to your computer, and if all you have is a 28.8Kbps modem, this can seem to take forever.

So let's look at the numbers behind it, and then illustrate it in terms of practical examples. The examples will give you a better understanding of what 'download speed' means in real terms, and what capacity you're likely to want, talking in terms of documents, streaming media, digital photographs and network gaming.

**Opening Web Pages vs File Downloads.**

The graphic elements of a web page – photos, buttons, backgrounds, drawings, and so forth – are all separate elements (or files) that are put together by your web browser software when you open a web page.

Each of these elements has a file size, although in the case of Web graphics, these file sizes are usually small. This is because Web file formats have been especially made to be very 'light-weight'; formats such as JPEG and GIF compress files so that they contain enough colour information to look good on screen, but not enough to print out at high definition. This minimizes their file size. A complex web page with many graphic elements may only be 55 kilobytes; however a single postcard-sized colour photo, if it is stored in a format suitable for high-quality printing, may be many megabytes, because it contains sufficient information for printing a large number of dots-per-inch on a high-resolution printer.

So while it is true that Web pages will appear more quickly on a broadband connection than via the fastest dial-up modems, the difference is not that great in simple web-browsing. It is in downloading files or streaming media types that the difference really starts to be seen.

With the amount of multimedia content now on the web, it is common to download audio or video files such as news footage, MP3 music files and the like. And because these kinds of files contain more information than simple text or graphic files, their file size is correspondingly larger.

## *Binary Systems Primer*

### How File Sizes are Measured[3]

The smallest unit of computer data is a single 'bit' which is a single digit in a binary number – either a ONE or a ZERO. Basically this represents the 'on' value or and 'off' value in a transistor, or on a magnetic storage device, or a tiny pit (or its absence) on a CD-ROM or DVD.

[3] http://computer.howstuffworks.com/bytes.htm/printable

Believe it or not, all of the vast array of information on the Internet, all the words, images, movies, ideas, news stories, songs, gossip, scandal and intrigue, all the diagrams, plans, ideas, websites, newsgroups, computer viruses and their cures, every program, every file, are ultimately just rows of ones and zeros, zipping around at the speed of light.

Mind you, there’s lots of ’em.

**Binary Systems**

So where common mathematics is counted in tens, binary data systems count in twos, using only two numerals, namely zero and one. A single binary digit is called a ‘bit’, which corresponds to a ‘digit’ in a decimal system.

Where decimal systems are based on powers of ten: $10^0$ (=1), $10^1$ (=10), $10^2$ (=100) and so on, binary systems are based on powers of two: $2^0$ (=1), $2^1$ (=2), $2^2$ (=4 ) $2^3$ (=8) $2^4$ (=16) and so on.

Powers of two ascend as follows: 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, and so on, which is a familiar sequence of numbers for computer users, as it is the standard measure, in megabytes, of the capacity of RAM chips, among other things.

The binary values for the numbers 1 – 15:

| Binary= | Decimal | Binary= | Decimal |
|---|---|---|---|
| 0000 = | 0 | 1000 = | 8 |
| 0001 = | 1 | 1001 = | 9 |
| 0010 = | 2 | 1010 = | 10 |
| 0011 = | 3 | 1011 = | 11 |
| 0100 = | 4 | 1100 = | 12 |
| 0101 = | 5 | 1101 = | 13 |
| 0110 = | 6 | 1110 = | 14 |
| 0111 = | 7 | 1111 = | 15 |

**Binary Code and Morse Code**

Binary code can be compared to Morse code, which is also binary.  Why? Because it is also based on on/off values: DOT or DASH.

SOS in Morse code is . . .  ---  . . .

In binary code, SOS is 010100110100111101010011.

So you could think of computers as devices which store and send everything by something similar of morse code, very quickly. But unlike the old Morse system, they also translate the code back into words, numbers, images and pictures.

**Bits and Bytes**

The smallest coherent unit of binary data is eight **bits**, or a decimal number comprising eight characters, which makes a **byte**.  It just so happens that the range of 8-bit numbers, or bytes, starting at 1 and ending at 11111111 (= decimal 255) are sufficient to describe a set of 255 characters encompassing the alphabet and the set of numbers and symbols used to depict basic arithmetic, as well as characters such as ‘&’ and ‘@’, accents, and so on.

Hence the common description of a byte: it is just enough data to describe a letter.

Data storage is generally measured in Bytes/Kilobytes/Megabytes as follows

| Name | Abbr | No of Bits | Size in Bytes |
|---|---|---|---|
| Kilobyte | kB | $2^{10}$ | 1,024 |
| Megabyte | mB | $2^{20}$ | 1,048,576 |
| Gigabyte | gB | $2^{30}$ | 1,073,741,824 |
| Terabyte | tB | $2^{40}$ | 1,099,511,627,776 |

And so on. Most computer users will never need to deal with a bigger unit of measurement than a gigabyte although office IT managers might nowadays need to consider terabytes when planning for server or back-up capacity.

It is likely that your computer's hard disk capacity is measured in Gigabytes (unless it was made before the mid-nineties), and that most of your files are measured in KB or MB (although uncompressed video and sound files can be GB).

**Kilobytes versus Kilobits**

Now one of the GENUINELY CONFUSING THINGS about data speed is that data transfer rates are usually referred to terms of KILOBITS and MEGABITS, not KILOBYTES and MEGABYTES, per second

Kilobits-per-sec is abbreviated as K**b**ps – note the small '**b**', while Kilobyte is abbreviated as kB or KB, with a capital B.

You may be tempted to ask, why this differentiation? Are those computer people just trying to bamboozle us again? The answer is that while data going across a communications link is logically measured in bits-per-second, data sitting on a storage devices is logically measured in kilobytes and megabytes. What this means is that data flowing through a modem at 256 Kbps is moving at **about** 30KBps. Later on we'll look at examples.

**What about 'Baud'?**

If you have been around computers for a while will remember the term 'BAUD' as a descriptor of modem data speed. On every phone line, the carrier signal is characterized by the number of signal intervals, or pulses, that are transmitted per second. Each pulse is called a 'baud'. Early modems were able to transmit one bit per baud, so at this stage the terms were regarded as interchangeable. However as modem technology progressed, a lot more bits were able to be carried per each pulse, or baud, and the term 'baud' fell into disuse. Nowadays modem speed is nearly always expressed in terms of kilobits-per-second.

**Other factors**

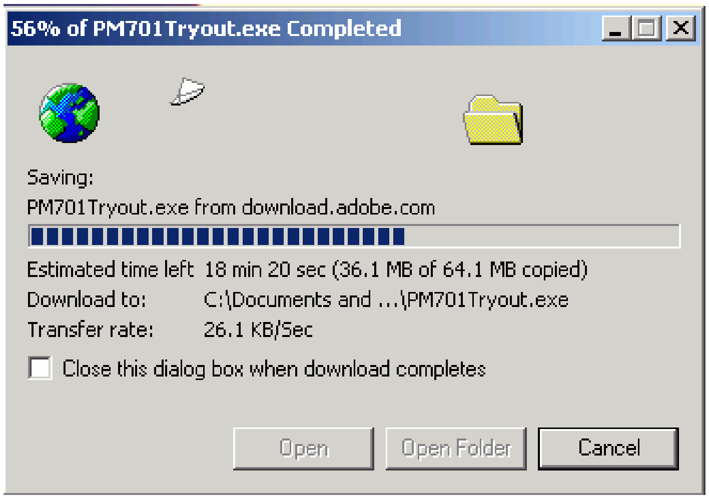
While we have seen that 8 bits makes the one byte necessary to describe a single character, in modem transmission of data an additional two bits of data are required to mark the start and end of a character (hence, start bit and stop bit); so transmitting a character across a modem takes 10 bits, not 8. Ergo, 56,000 bits per second does not exactly amount to the number of kilobytes you might expect, due to the additional data that needs to be sent for each character.

Finally, any kind of network connection is subject to an indefinite number of variables which can degrade performance, including amount of 'noise' on the line, condition of various routers and switches between the two points, and the performance of the originating web server.

The following table presents IDEAL download speeds without taking into account 'stop-bit disparity' and the fact that real network speeds are affected by line quality and other factors.

| **Connect Via** | **Kilobits/ sec** | **Bits/sec** | **Bytes/sec** | **KByte/M in** | **MBytes/M in** |
|---|---|---|---|---|---|
| Slow Dial-up | 14.4 | 14400 | 1800 | 105 | 0.10 |
| Medium Dial-up | 28.8 | 28800 | 3600 | 211 | 0.21 |
| Fast Dial-up | 56.6 | 53000 | 6625 | 388 | 0.38 |
| Basic ADSL | 128 | 128000 | 16000 | 938 | 0.92 |
| Mid-range ADSL | 256 | 256000 | 32000 | 1875 | 1.83 |
| Fast ADSL | 512 | 512000 | 64000 | 3750 | 3.66 |

If you're a Win XP user, you may notice that when you start a large file download, Windows XP depicts the download speed in KILOBYTES PER SEC (see below). Conversely, your modem's syslog page, be it ADSL or Dial-up, is likely to depict this information in KILOBITS PER SEC.  It is easy to confuse the two, but the dialog box below actually depicts a transfer rate of around 208 kilobits (=208,000 bits) per sec, coming across a 256 Kilobits/sec connection.



56% of PM701Tryout.exe Completed

Saving:

PM701Tryout.exe from download.adobe.com

Estimated time left 18 min 20 sec (36.1 MB of 64.1 MB copied)

Download to: C:\Documents and ...\PM701Tryout.exe

Transfer rate: 26.1 KB/Sec

☐ Close this dialog box when download completes

## Real World File Sizes

To get an idea of file sizes, look at the **approximations** in the following table:

Unformatted e-mail message 600 words = 48,000 bits = 6,000 bytes = **5.85Kb**.

Formatted word document 600 words = 208,000 bits = 26,000 bytes = **25.4 Kb**

JPEG Colour photo, 23 x 16.5cm, 92 dots-per-inch resolution = 736,000 bits = 92,000 bytes = **90Kb**

JPEG Colour photo, same area, 300 dot-per-inch resolution = 418,000 bytes = **408Kb**

TIFF Colour photo, same area, uncompressed, high res = 100,800,000 bits = 12,600,000 bytes = 12,304Kb = **12.30Mb**

3.5 minute song, stereo, CD-quality, uncompressed = 296,000,000 bits = 37,000,000 bytes = 36,132Kb = **36Mb**

3.5 minute song, stereo, MP3 compression = 27,200,000 bits = 3,400,000 bytes = 3,320Kb = approx **3.3Mb**.

Now let's consider the **THEORETICAL** time it would take to download these files at various network speeds ('theoretical' because unpredictable factors affect the outcome especially at higher speeds.) [4]

| | Plain text message | WP doc | Colour Pic 92 dpi res | Colour Pic 300 dpi | Colour pic Hi Res | Digital Audio 'Raw' | Digital Audio MP3 |
|---|---|---|---|---|---|---|---|
| File Size | 5.85Kb | 25.4Kb | 90Kb | 408Kb | 12.30Mb | 36 Mb | 3.3Mb |
| Slow Modem 28.8kbps | 1 second | 7 Sec | 3 sec | 2m 7s | 1hr 4m 3s | 3h 7m 30s | 17m 11s |
| Fast modem 56 kbps | ⚡ | 3 sec | 1.8 sec | 56s | 28m 35s | 1hr 23m 42s | 7m 40s |
| Broadband 128Kbps | ⚡ | ⚡ | 5 sec | 24s | 12m 30s | 36m 37s | 3m 20s |
| Broadband 640 Kbps | ⚡ | . ⚡ | ⚡ | 5s | 2m 33s | 7m 30s | 41s |

## File Compression

You will notice in the above table the difference between the 'raw' audio data file at 36Mb and its MP3 version which is less than a tenth of the size.

This difference is a result of 'file compression'. Basically, file compression works by analysing the data that comprises the file and then re-coding it in the form of a pattern which eliminates a lot of the repetitive information which typically occurs in files.

Ordinary application data, such as word-processor, spreadsheet and database files, can easily be compressed using commonly-available compression utilities. Such compression methods are based on the principle that a large proportion of the information contained in these kinds of files is alpha-numeric, which is very easy to represent by various kinds of shorthand.

Compression technologies that allow files to be restored to their exact, pre-compression condition are called 'lossless' in that no data is lost in the file compression process. Lossless compression is generally used where no variation in the file content or format is allowed, for example spreadsheet or text documents and in technical drawings.

Examples of lossless compression utilities and file formats include WinZip and, on the Macintosh platform, Stuffit. PNG is a lossless file compression format for graphics while the older JPEG and GIF formats both sacrifice image data and will be mentioned below.

Multimedia, image and graphic content make use of more sophisticated compression techniques which reduce the amount of information in the file (hence

[4] http://www.martindalecenter.com/AATimeCalc.html

the term 'lossy compression') by eliminating or simplifying elements that are redundant, inaudible or imperceptible.

The example in the table is MP3 which has become hugely popular for downloadable music. The 'MP3' acronym is derived from 'Motion Picture Experts Group' (i.e. 'MPEG') which developed the technology, and the numeral refers to the 'audio' layer of a file (=layer 3). MP3 provides acceptable playback quality while reducing the file size by an order of magnitude. Apple's iTunes is, however, built around another compression format, namely AAC[5], or Advanced Audio Coding, which offers a greater frequency range and other benefits.

Lossy graphic compression formats include JPEG and GIF, for moving pictures, MPEG1, 2 and 3 and Motion JPEG.

**Using File Compression**

Current versions of Windows have WinZip built in. Web browsers provide translators to automatically interpret common image compression formats such as JPEG or GIF. In other cases you will need to acquire the appropriate software or hardware to compress or decompress the file; for example, the iTunes data format can only be interpreted by an Apple iPod music player; and some kinds of MPEG video require a hardware device to be installed in your computer.

## Streaming Media

'Streaming media' formats are used to provide 'just-in-time' multimedia information, such as Real Audio or Windows Media data streams. Streaming enables the file to be played as it arrives, instead of being completely downloaded prior to playback. Many large commercial news services offer streaming media news bulletins on their sites.

Streaming media offers one of the most compelling arguments for a broadband connection. Across dial-up connections, many streaming media files run intermittently and in very small windows at low quality. While streaming media on a high-bandwidth link has still got a way to go before it looks like television, if it is produced properly it is very effective – check out the showcase on Macromedia.com for examples.

When a streaming media file is downloaded, you will frequently see that the file is being 'buffered', meaning it is being stored on a cache on your system prior to playback. This technique facilitates uninterrupted playback by creating a buffer of data while still downloading the file from the source.

Common streaming media formats include

- Windows Media Player
- Real Networks: Real Player
- Apple Computer: QuickTime
- Open-source : Ogg Vobis.

Common protocols include RTPS, for Real-Time Streaming Protocol, which allows use of Stop-Pause-Replay controls as well as the HTTP protocol.

The compression algorithms used for streaming media are known as 'codecs' for compression/decompression. The data is compressed for transmission across the Internet, and then decompressed on receive.

[5] http://en.wikipedia.org/wiki/Advanced_Audio_Coding

# The NetComm DSL Primer

### DSL: Exploiting Unused Capacity on Telephone Lines

Standard telephone services are carried on a pair of copper wires coming into a house or office. DSL [=Digital Subscriber Line] technology uses the spare capacity on these phone lines to move data at much higher rates than that offered by older-style analogue modems.

Voice (and fax) traffic only occupies a small segment of the lower-frequency spectrum available on copper wires, from 0 to 3,400 hertz [= cycles per second]. DSL uses sophisticated algorithms to package and transmit data across the same wires at much higher frequencies, typically using frequencies above 4Khz. This allows the same wire to carry much more data than it was originally intended for, while allowing the co-existence of voice, fax and data services on the same line, provided filters are used to separate the signals.

### ADSL Filters

ADSL Filters operate by removing signals of all frequencies above 4KHz from the service. If you connect a telephone to an unfiltered ADSL line, you will hear a lot of electrical interference from the high-wavelength data transmissions traveling along the wire and the telephone may not even function. Splitters and filters isolate the data-service frequencies and the voice frequencies from each other.

### Types of DSL Connection

### ADSL

Asymmetric DSL [=ADSL] allows much faster downloads than uploads because it sacrifices up-channel bandwidth in favor of the down channel. For most users, however, this asymmetry is not a problem, as they are accustomed to downloading a lot more information than they send.

ADSL is not available everywhere, however. The proximate telephone exchange must be equipped for ADSL. Also the effective wire-length (or ‘local loop’) from the telephone exchange to the service recipient is currently deemed by Telstra to be 4,000m (bearing in mind that wire-length may be different to geographical distance).

ADSL data transmission speeds start at 256Kbps download/64Kbps upload, up to 1.5Mbps/512K. Note that theoretical capacity exists for 8Mbps download and 800Kbps upload via ADSL but due to equipment constraints this speed is not generally obtainable. ADSL is rate adaptive and will automatically detect the maximum up and down speeds possible for a given line quality and length.

### ADSL2 and ADSL2+

Described by standards G.992.3 and G.992.4, and approved by the International Telecommunications Union, these new technologies will soon start to become available. ADSL2 improves on the original ADSL implementation by increasing download rates from 8 to 12Mbps, as well as by extending the local-loop length by up to 200 metres. ADSL2 also offers improved diagnostics, better power management with stand-by mode, greater interoperability, faster start-up, and support for packet-based services (e.g. Ethernet). ADSL2+, a little further out, is based on ADSL2 and doubles the maximum frequency used for downstream data transmission from 1.1 MHz to 2.2 MHz. As a result, downstream data rates are increased to up to 24 Mbps on phone lines as long as 3,000 feet, and 20 Mbps on lines as long as 5,000 feet.

### SHDSL

Symmetric High Bit-rate DSL [=SHDSL] offers the same data transmission rate for both upload and download: 2.3MBps for a two-wire, or 4.6Mbps for a four-wire model. [A normal telephone service comprises two wires, while two telephone lines are needed to provide a four-wire connection.]

SHDSL conforms to the ITU standard G.991.2 and is also referred to as G.shdsl.

SHDSL uses the full spectrum of the line and therefore needs a dedicated line and cannot share a line being used for voice/fax traffic. It supports a greater local-loop distance than does ADSL, and is able to transmit data at 1.2Mbps over 6,100 meters wire-length from the exchange. SHDSL is rate adaptive and will automatically detect the maximum up and down speeds possible for a given line quality and length.

ADSL and SHDSL are the most common types of DSL service in the Australian market at time of writing, but for completion several other variants will be mentioned.

### VDSL

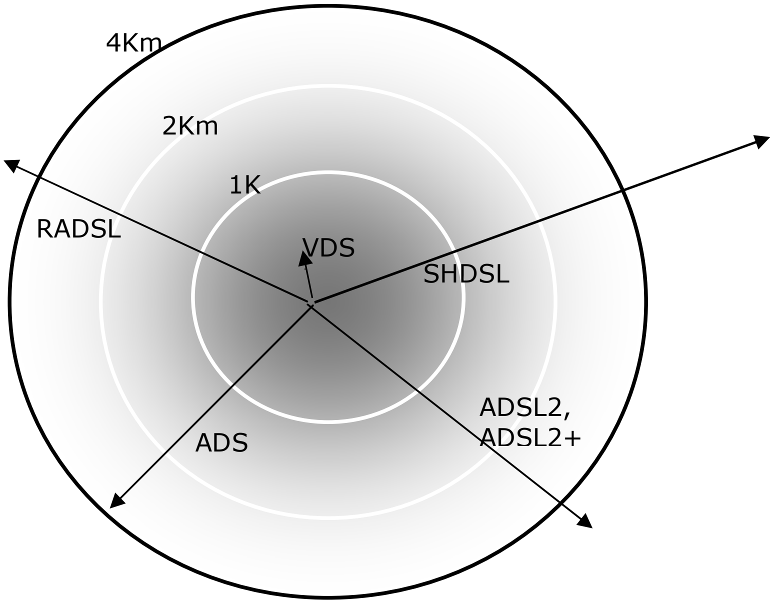
Standing for 'Very-high Bit Rate DSL', VDSL is an emerging technology that is intended to provide a high-speed link between homes and a nearby fibre-optic cable. The requirement for 'nearby cable' means that deployment is probably not imminent for most users. VDSL offers up to 52Mbps over short distances and is being offered in some areas in the US.

### RADSL

'Rate Adaptive DSL' is an asymmetric technology that allows the modem to adapt the rate of transmission to 7Mbps downstream and 1MBps upstream, depending on variables such as file size and line quality. RADSL standards are certified but the protocol is yet to be widely adopted.

Other standards are IDSL, for ISDN over DSL; SDSL, an emerging enhancement to existing SHDSL; and HDSL, an older version of Symmetric DSL.

Fig 1: Approximate Maximum Local Loop Length for DSL Technologies



## Table 1: Summary of DSL Standards June 2004

| Type | Direction | Upload Max | Download Max* | # Lines | Phone Support |
|---|---|---|---|---|---|
| ADSL | Asymmetric | 800 Kbps | 8 Mbps | 1 | Yes |
| ADSL2 | Asymmetric | 800 Kbps | 12Mbps | 1 | Yes |
| SHDSL | Symmetric | 2.3/4.6 Mbps | 2.3/4.6 Mbps | 1, 2 | No |
| HDSL | Symmetric | 1.54 Mbps | 1.54 Mbps | 2 | No |
| IDSL | Symmetric | 144 Kbps | 144 Kbps | 1 | No |
| RADSL | Either | 1 Mbps | 7 Mbps | 1 | Yes |
| VDSL | Asymmetric | 16 Mbps | 52 Mbps | 1 | Yes |

* Notes: The 'maximum download' rates for ADSL are not often reachable due to various system constraints with most consumer-grade services being advertised at 1.5Mbps download max.

**NetComm DSL Solutions**

NetComm provides a range of state-of-the-art ADSL and SHDSL modem/router/gateways.

The **DSL modem** connects to the phone line and receives data from the DSL Access Multiplexer [DSLAM] at the exchange for transmission via the customer premises equipment [CPE]. The **router** component also performs routing functions, namely routing network traffic between two network segments [e.g. the Internet and your LAN]. The models listed here are also **gateways** in that they maintain the DSL connection even while no connected computer is powered up.

## Table 2: NetComm DSL Solutions Matrix June 2004

| | Modem/Router | Load Balancing/ Line Aggregation |
|---|---|---|
| ADSL | NB1200<br>NB1300, 1300+4, 1300W<br>NB5580, 5580W | |
| SHDSL | NB702 2-Wire<br>NB704 4-Wire | NB740 |
| ADSL2 [+] | TBC | |

# Chapter 5: Introduction to Wireless Networking

Wireless networking uses radio transmission instead of wires to provide high-speed data access to computers and networks. Built around the 802.11 series of standards and implemented by most mainstream IT vendors, it has taken off in a big way since 1999.

Wireless networking provides instantly-available public access to the Internet and networked computer resources in venues such as diverse as airports, cafes and libraries, and is increasingly used to provide convenient and scalable information access on University campuses, and in health-care, retail, and manufacturing environments.

Wireless networking can be easily and seamlessly combined with advanced software applications embedded in, for example, PDAs, Tablet PCs and personal computers to facilitate all sorts of interactive and computer-assisted tasks and activities.

For home users and small office environments this medium has the attraction of requiring no physical wiring, as well as the convenience of being able to sit wherever you like. Read the daily papers online at breakfast, answer your emails on the back deck, download music at an airport lounge – you can connect from virtually anywhere.

The sky really is the limit with wireless networking. Just as TV and radio have become taken-for-granted features of the modern environment, so too will wireless networking and ubiquitous information services become part of the social fabric in the 21st century. Indeed, they already have.

## Wireless Networking History

Radio data transmission was first used by the U.S. Army and its allies during the Second World War. This made use of advanced encryption, a science which was dramatically advanced during the war.

Frequency-Hopping Spread Spectrum technology was also developed by military research and was first used during the Cuban Missile crisis in 1962. In FHSS, the signal 'hops' between a number of adjacent frequencies to prevent jamming and eavesdropping, a technique which has become fundamental to wireless networking since then.

### AlohaNET

The next milestone in the history of wireless was the creation of **AlohaNET**. Norm Abrahamson got interested in the Internet because he wanted to surf - not 'surf the Web', which didn't even exist then, but surf in Hawaii! An engineer, he got a job at the Faculty of Engineering at the University of Hawaii and in 1970 created a wireless network which linked three islands of the Hawaii atoll to a central computer on Oahu Island. AlohaNET was also a training ground for Bob Metcalfe, who went on to develop Ethernet, and some of the principles behind AlohaNET were adapted for use in the first Ethernet networks.

### Allocation of the ISM Band

Then in 1985 the US Federal Communications Commission [FCC] designated a small band of the available radio spectrum to unlicensed use for Industrial, Scientific and Medical purposes, thereafter known as the **ISM band**. Part of the specification was that (1) transmissions had to be limited to one watt and (2) spread spectrum transmission technology had to be used. This was to enable neighbouring installations to co-exist without interference and was one of the far-sighted decisions which has enabled wireless networking to reach the high level of usefulness that it has today.

### Emergence of Standards

However the missing part of the puzzle was the development of standards, as at this stage vendors servicing the fledgling ISM-band industry were building systems to different, and incompatible, specifications. It was not until the release of IEEE[6] 802.11 in 1997 that a standard was developed allowing inter-operability between products made by different vendors.

### Apple's Airport

The first commercially significant implementation of wireless networking was the introduction of Apple's 'Airport' wireless networking in 1998. Built around the then-to-be-ratified 802.11b standard, Airport represented yet another example of Apple Computer's oft-demonstrated ability to be first to market with a technology which was soon to be wanted by everyone (more info in breakout, below.)

### Wireless Hotspots

Since 2000 there has been a frenzy of activity in deployment of 'Wireless Hotspots', or public access points, originally centred on airline terminals and trendy inner-city coffee shops but quickly spreading to embrace 'hot zones' covering whole neighbourhoods, districts, university campuses and industrial estates[7]. Lufthansa was first among airlines offering hotspot access in the air, commencing in January 2003. In Europe, IDC reported that the number of Hotspots increased by 327 percent during 2002, from 269 locations at the end of

[6] Institute of Electrical and Electronic Engineers - the organisation which sets standards for networking interoperability.

[7] http://www.wi-fiplanet.com/special/index.php/11051

2001 to 1,150 at the end of 2002.[8] Development continues to explode, with large- and small-scale installations appearing all over the world[9].

## *Wireless Standards*

### The Emergence of IEEE 802.11

Work on standardisation is ongoing, conducted mainly by the Institute of Electrical and Electronics Engineers (IEEE), the Internet Engineering Task Force (IETF), the Wi-Fi Alliance, and the International Telecommunication Union (ITU). Further capabilities and refinements are continually being devised, resulting in a thriving industry supported by vendors producing interoperable wireless data devices.

Let's look more closely at Wireless Standards to get more of an idea of what they offer and how they differ.

The IEEE 802.11 standard drafted in 1997 described how the methods for encoding and transmitting data between computers using radio signals had to work.

Originally, the standard defined three types of radio connection, specifically Direct Sequence Spread Spectrum [=DSSS], Frequency Hopping Spread Spectrum [=FHSS] and Infra-Red [=IR], supporting data rates of one to two megabits per second.

Further refinement of the specification in 1999 resulted in the **802.11b** and **802.11a** versions of the standards.

**802.11b** provides (theoretical) data rates of 5.5 and 11Mbps in DSSS mode and became the basis of Apple's Airport implementation, which actually shipped before the standard was finally ratified in 1999. ('Theoretical' is always the case with maxima quoted in wireless networking, as the actual speed is subject to a number of constraints including the fact that a large percentage of the bandwidth is required for error-checking and protocol management to ensure that the data flow is not corrupted.)

802.11b became hugely successful, and has been the basis for much of the growth that has occurred since, despite some advantages inherent in the 802.11a standard.

### 802.11g

Subsequent refinements to the 'b' standard resulted in **802.11g** which transmits on the same wavelength as 'b' but which offers higher download speeds up to a theoretical maximum of 54 Mbps. This improvement was mainly due to the adoption of a technology called Orthogonal Frequency Division Multiplexing (about which more elsewhere).

The transition from 802.11b to 802.11g was made over the period from 2002-2003 and was again anticipated by Apple with the release of Airport Extreme, although in this case several other network product vendors shipped 802.11g units before Apple, and indeed even before the end of 2002.[10] There was

[8] IDC, quoted in http://www.wi-fiplanet.com/news/article.php/1581141

[9] Including some far-flung corners; see http://nepalwireless.net/index.html for how wireless is benefiting Nepalese yak farmers

[10] http://www.oreillynet.com/pub/a/wireless/2003/01/23/80211g.html?page=2

commercial risk in so doing as the standard was not ratified until mid-2003[11], meaning that vendors were assembling products on the basis of draft specifications. However, in the event, the draft specifications stood, and 802.11g is now setting the pace, not least because it is backwards compatible with the 802.11b standard.

### What About 802.11a?

As said above, 802.11a does offer technical superiority in the some respects.

First, 802.11a broadcasts on a different frequency, namely 5MHz as opposed to the 2.4MHz used for the other versions. Higher frequency allows higher data transmission rates (although at a shorter range).

Secondly, this frequency band in not nearly as heavily used as the 2.4MHz band, and is therefore less likely to receive interference from other devices.

Third, the .11a standard allows for more non-overlapping channels due to the wider spread of available spectrum in the 5MHz range, which means that it can support a higher number of clients in the same location without performance degradation or interference.

Finally, even though the .11g and .11a standards are theoretically capable of the same maximum download rates of 54Mbps, .11g networks can suffer performance slowdowns when hosting .11b clients, as the network adapts to deal with the slower systems.[12] Testing also indicates that overall network performance is faster with the 802.11a standard.

All of this said, however, at time of writing devices based on the 802.11b/g standard comprise the vast majority of installed Wireless LANs and hotspots, and despite 802.11a's advantages it is already being dismissed by some observers as the Betamax of broadband. The truth of this remains to be seen, and many vendors are offering hybrid wireless access devices which can work with all three standards. In terms of market share, however, the 802.11b and g versions are streets ahead at this time.

### 802.11i and beyond

802.11i, ratified by the IEEE on June 25th, is mainly concerned with network security, and is designed to rectify the inherent security weaknesses detected in WEP (=Wireless Equivalent Privacy). The interim WPA (=Wi-Fi Protected Access) solution was brought out in June 2002 in advance of the 802.11i ratification by the Wi-Fi alliance in response to commercial pressure, amounted to a sub-set of the features ratified by 802.11i.

Main features of 802.11i are:

**TKIP** (Pronounced Tee-Kip), or Temporal Key Integrity Protocol, is an encryption method designed specifically to address the shortcomings of WEP. Basically, one of those shortcomings was the repetition of encryption keys over a potentially short time cycle; if enough data was intercepted from a WEP-protected network over a period of hours, the repetition of the key would allow it to be deduced. TKIP adds a Temporal Key which is changed every ten thousand packets (i.e. very

[11] On June 12th 2003 to be precise, see http://standards.ieee.org/announcements/80211gfinal.html

[12] http://www.nwfusion.com/techinsider/2002/0520wlan/0520feat2.html and http://www.nwfusion.com/columnists/2004/0524wizards.html for comparisons of the 3 standards.

frequently), combined with the client MAC address and a 16-octet initialization vector, making the key practically impossible to deduce.

802.11i also includes provision for AES (=Advanced Encryption Standard) which may however require hardware upgrades in a many networks, and enables incorporation of two-factor authentication systems such as smart-cards.

802.11i incorporates RADIUS authentication (=**R**emote **A**uthentication **Di**al-in **U**ser **S**ervice), an industry-standard access-control server technology.

Work is ongoing in the 802.11 group, with further workgroups developing standardisation on issues such as increasing data through-put, fast hand-off between wireless cells, and self-healing network topologies.

### WiMAX and WAP

Grouped around IEEE standards numbered 802.16, WiMAX is designed to service whole neighbourhoods rather than small 'hotspots', so it is not a local-area network technology. WAP stands for Wireless Application Protocol and is used to provide Internet services to mobile telephones and PDAs. Both protocols are outside the scope of this paper which is concerned with Wireless LAN.

### What about Super G?

Super G or Turbo G is not a new standard, but a proprietary method of boosting data transmission rates across 802.11g networks using frame-bursting and dual channel techniques. Produced by Atheros, the technology is available in networks composed of equipment built around the Atheros chip-set. NetComm's initial Super G devices are about to debut.

# Intel Centrino

Intel has acknowledged the importance of wireless networking by the introduction of the Centrino product line which provides motherboard-level support for wireless notebook connection.  Centrino consists of a combination of 802.11a/b/g-compliant chips and power-management and networking utility.

Current product range includes:

Intel® PRO/Wireless 2100 Network Connection (Single band 802.11b), Intel® PRO/Wireless 2100A Network Connection (Dual band 802.11a/b), Intel® PRO/Wireless 2200BG Network Connection (Dual mode 802.11b/g)

Performance:  Typical indoor range of 100 ft (30 m) @ 11 Mbps / 300 ft (90 m) @ 1 Mbps for 802.11b

Other Features:

Intel® Wireless Coexistence System support enables reduced interference between Intel PRO/Wireless and certain Bluetooth devices.

For systems designed with two antennas, real-time antenna selection enables optimized WLAN performance.

Real-time temperature calibration dynamically optimizes wireless performance by adjusting output power to temperature changes for increased throughput & range with 802.11a radio

For systems designed with two antennas, real-time antenna selection enables optimized WLAN performance.

Intel® **PROSet** software allows for multiple profile setup and automatic switching between profiles for simplified wireless access as you move between different access points. Allows for ease of setting available security options. Supports Cisco, Check Point Software Technologies, Microsoft and Intel VPN connections.

Power Management: **Intelligent Scanning Technology** reduces power by controlling the frequency of scanning for access points. User selectable feature with five different power states, which allows the user to make their own power vs. performance choices when in battery mode.

More info at:  http://www.intel.com/products/mobiletechnology/

## Wireless Networking Terminology

There are some terms that need to be understood to make sense out of wireless networking. Many of these will be explained in much greater depth later in the document.

### Ad Hoc Mode

It only takes two computers equipped with wireless cards to create what is known as an 'Ad Hoc Network'. In this configuration, files can be shared among personal computers using capabilities provided by the operating system. Ad Hoc networks can accommodate small workgroups however cannot provide access to network resources nor to the Internet.

### Infrastructure Mode

When Wireless Access Points ['AP' – see below] are introduced, this is referred to as an infrastructure-mode WLAN. Infrastructure-based WLANs provide the full suite of networked services including internet access, file and printer sharing, and more.

### WLAN

Wireless Local Area Network. When discussing workgroup networking, this is what is usually being referred to.

### SSID

SSID stands for Service Set Identifier. A Service Set is defined as a group of Stations [=PCs with wireless cards] located within sufficient proximity for all to communicate. So, the Service Set Identifier is your wireless network name. As it is possible to be receiving signals from a number of WLANs in one place, you may need to enter the SSID to access the right **Security tip**: in setting up a WLAN, use a nondescript SSID, as the name will be broadcast by radio: if you call your network by a literal or obvious name it may invite attention from the wrong listeners [i.e. when in a hostile neighbourhood - 'dress down and avoid eye-contact'].

### WEP Key

WEP stands for Wired-Equivalent Privacy, or security equal to that which is obtainable on a wired network. An encryption system which exists in two variations, 64-bit or 128-bit, WEP has been subject to criticism on the basis of studies which showed that WEP schemes are inherently insecure[13]. However it should be noted that most unauthorised access occurs because of lax security practises, not because a network has been hacked; despite its flaws, WEP is still too hard for an average user to crack. WEP is, in any case, in the process of being upgraded to the more secure WPA standard which promises to overcome the weaknesses of WEP.

From a user viewpoint, what you will often need to know is the **WEP key** which is, to all intents and purposes, a pass code you will need to enter to access a wireless network.

[13] http://www.cnn.com/2001/TECH/ptech/08/10/wireless.hack/index.html

### Encryption Type

A related function, users may need to specify the type of security that is operating on the network they wish to join, namely 64 or 128-bit WEP or WPA, prior to entering the pass code.

### Wireless Access Point

The Wireless Access Point [=AP] is the basis of the infrastructure-based WLAN. The Access Point functions in a similar way to a Mobile Phone base station. The AP **bridges** wired and wireless networks; you can easily connect an AP to a wired router, providing Wireless access in addition to existing wired LAN access. In this situation, the AP sends, buffers and receives data between the WLAN and the whole wired world.

### Wireless Channels

Wireless Channels are created as subdivisions of the frequency allocated to Wireless LAN connection.[14] IN the U.S. 802.11b/g supports 11 channels, in Australia, 13 channels. In an ad-hoc network, channel selection needs to be entered on each client computer to enable them to communicate, while in an Infrastructure-based network, channel management is automated by the Access Point(s).

### Wireless Roaming

Roaming allows client computers to roam between Access Points, meaning that if they move out of range of one AP, they can seamlessly connect with the next. Roaming is a function that is enabled by the 802.11 standards and relies on using the SSID to integrate different APs into one network.

### Fallback

802.11 provides ‘graceful degradation’ capacity, which means that if the signal strength begins to fail or interference occurs, the network will sacrifice throughput for reliability, on the basis that it is better to have a slow connection than to have the connection fail.

### AP Positioning

Environmental and architectural factors often influence wireless network connectivity. Microwave ovens and cordless telephones are both frequent sources of interference for WLANs. As well as this the proximity of large metal objects (e.g. oil storage tanks), double-brick or concrete walls, and other barriers can affect network reliability. The result is in some installations ‘dead zones’ may be encountered where client machines loose contact with the network; often this can only be solved by moving the Access Points around until a better location is discovered by trial and error.

[14] Although 802.11a offers more channels than 802.11b/g because it has a wider frequency range, we’ll concentrate on 11b/g because it is more widely used.

## *Summary Comparison of Wireless Standards*

| Standard Name | **802.11b** |
|---|---|
| Prevalence: | Largest installed base[15] |
| Operating Frequency: | 2.4 - 2.4835GHz ISM band |
| Channels: | 11 channels - 3 non-overlapping, each channel 22 MHz centred at 5 MHz intervals beginning at 2.412 GHz and ending at 2.462 GHz |
| Data Rates: | 1, 2, 5.5, and 11 Mbps<br>avg actual throughput of 4.5Mpbs. |
| Capacity: | 32 users per access point |
| Pro & Con: | Longer range than 802.11a, lesser network density |
| Transmission Method: | DSSS (Direct Sequence Spread Spectrum) |
| **Standard Name** | 802.11g |
| Prevalence: | Minor but expected to grow quickly |
| Frequency, Channels | As above |
| Data Rates: | Increased to ideal 54Mbps<br>Avg actual throughput 7-16Mbps |
| Pro & Con: | As above but much higher throughput. Compatible with 802.11b installations. |
| Transmission Method: | OFDM (Orthogonal Frequency Division Multiplexing) |
| **Standard Name** | 802.11a |
| Prevalence: | Smaller installed base |
| Operating Frequency | 5GHz UNII (Unlicensed National Information Infrastructure) band |
| Channels: | 12 non-overlapping channels, each 20 MHz wide, centred at 20 MHz intervals (beginning at 5.180 GHz and ending at 5.320 GHz for the upper and middle U-NII bands, beginning at 5.745 GHz and ending at 5.805 GHz for the |

[15] 802.11b accounts for about 95 percent of the units shipped and about 98 percent of the total value of the market in 2002. "802.11a-only sales are really tailing off. People are waiting now for dual-mode systems." – 2002 figures, quoted in http://www.unstrung.com/document.asp?doc_id=28303

| | |
|---|---|
| | upper U-NII band). |
| Data Rates: | Mandatory data rates of 6, 12, and 24 Mbps and Optional data rates of 9, 18, 36, and 54 Mbps<br>Avg actual throughput 7-16Mbps |
| Pro & Con: | Greater user density, but shorter range and less ability to penetrate barriers |
| Capacity | 64 users per access point, and more access points can be co-located due to greater available bandwidth |
| Transmission Method: | OFDM (Orthogonal Frequency Division Multiplexing) |

# Intro to Voice Over IP

VoIP is a method of sending digitally encrypted voice transmissions across the internet, bypassing the standard (PSTN) telephone network.

VoIP is a class of peer-to-peer application, which uses 'handshaking' to establish a link and a subsequent data-stream between peered IP devices, which may be IP phones, computers, PDAs, or other types of device.

VoIP call-setup and other signalling functions typically involve the exchange of relatively short message components comprising single packets rather than long packet sequences. Accordingly, VoIP signalling is usually implemented in UDP (User Datagram Protocol), rather than the more robust and 'heavier' TCP, though this is sometimes used in specialised VoIP applications. UDP contains less error-correction and verification information than does TCP thereby reducing 'overhead'.

Once data exchange – e.g. a conversation – is initiated, encoded audio is carried in the payloads of a sequence of short packets, each preceded by three protocol headers:

- **IP header** carrying the IP addresses of sending and receiving devices with a flag signifying the data type
- **UDP header** carrying the source and destination logical ports
- **RTP (Realtime Transport Protocol) header** carrying packet sequence and timing code.

RTP is fundamental to VoIP as data packets must be translated in the right sequence to produce intelligible speech. In fact VoIP can be thought of as a type of streaming media, similar to QuickTime or Real Audio, using RTP to ensure the voice data does not become garbled or disordered during transmission.

In typical VoIP bitstreams, the IP and UTP headers may contain more data than the audio payload itself.

VoIP media packets are transported across the intervening IP network the same way as any other data traffic. At the remote device, they are re-sequenced, decoded and the payload data played back, reproducing the input audio signal as a copy of the original sound.

It is worth reflecting that VoIP is a practical possibility because of the sheer processing power of modern computer chips; the microchips powering computers 10 years ago would not have been able to cope with the load even if these processing techniques existed then.

**Digitization and Encoding**

Voice over IP starts with analogue data, specifically a pattern of frequencies initially produced by the larynx, propagated through the air, and transformed by microphone into an analogous pattern of electrical voltage fluctuations. So far, this is the same process as is performed by any telephone.

The signal is then converted to a sequence of binary numbers by an analogue-to-digital converter circuit, which samples line voltage at a high rate of speed, usually around 8,000 times per second.

The digitized input signal is then pre-processed by algorithms which filter out background noise and enhance speech-band information. When a speakerphone or audio conferencing equipment is being used, additional pre-processing eliminates reverberation, echo, and other kinds of interference.

VoIP algorithms also recognise and 'bracket' periods of silence so as to preserve bandwidth by not transmitting the sound of silence. The complete absence of sound is replaced with a sound to emulate normal telephone background sound, called 'comfort noise'. VoIP software might also recognise touch-tone sounds and replace them with digital-equivalent values (used, for example, when navigating a touch-tone-driven phone menu.)

Having been pre-processed, the input sound is then transformed mathematically into a sequence of packet payloads, and usually compressed for greater efficiency. Compression is accomplished by a VoIP 'codec' (=coder/decoder)

### VoIP CODECs

Some VoIP codecs enable VoIP across only 8 or 5 Kbps of bandwidth. However the cost for minimising bandwidth requirements includes 'latency' or audible delay in transmission which is characteristic of VoIP conversations across dial-up connections.

Now that more bandwidth is generally available due to the prevalence of broadband and 100BaseT Ethernet, VoIP has standardised on several codecs that are rather less miserly with bandwidth.

The most common VoIP codec for high-bandwidth links is **G.711**, an international standard for encoding telephone audio data on an **64 kbps channel** although with the overheads required by packet header data this usually occupies more than 80 Kbps. This codec is therefore not suitable for use with ADSL 64Kpbs uplinks as it does not provide sufficient 'available talk bandwidth' especially if the line is providing other services simultaneously. Several lower-bandwidth alternatives include G.729A and G.723.1. Characteristics of these codecs are summarised in the table below:

| Standard | PPS Duplex | Raw Data Duplex | IP Data Duplex | IP + LAN Duplex | IP + VLAN Duplex |
|---|---|---|---|---|---|
| G.711 | 132 | 128K bits/s | 174K bit/s | 193K bit/s | 198K bit/s |
| G.729A | 200 | 16K bits/s | 80K bit/s | 108K bit/s | 115K bit/s |
| G.723.1 | 66 | 10.6 or 12.8K bits/s | 33K bit/s | 43K bit/s | 45K bit/s |

where 'PPS' = packets per second, 'duplex' indicates two-way transmission. Other columns reflect the data 'overhead' imposed by different layers of the network stack. Many other codecs are also in circulation; see links for references.

### Mechanics of VoIP Transmission

The regular PSTN network is nowadays largely digital, with voice data being transformed into digital format for transmission across ATM networks, and the resulting bitstream being synchronised for end-to-end transmission by network clocks. In so doing, the transmission path emulates the behaviour of the wire that would have connected the two ends of the conversation in an older-style network.

Sending a sequence of VoIP packets across an IP network is a different thing altogether. IP was designed NOT to require point-to-point connection; instead IP packets are routed in response to prevailing network conditions by any collection of a number of possible ‘network hops’. So a bunch of VoIP packets from the same data-stream might all arrive at their endpoint by different routes, having to take a different number of ‘hops’. One effect of this is ‘network latency’, or lag. In other cases packets may be dropped, or arrive out of sequence, or at widely-varying time intervals, creating **‘jitter’.**

Now for ordinary, non-real-time data such as documents or email, TCP provides for error correction, packet sequencing, and re-transmission to assure correct transmission. However TCP/IP's approach to connection management is inappropriate for voice data which of necessity is **isochronous** (or time-dependent) between sender and receiver. Requesting retransmission of dropped packets would cause breaks in playback, or require the buffering of large amounts of data which would also ruin the conversation by requiring you to wait while speech was ‘buffered’ into a data pool (‘hold on for a sec…the telephone is thinking….’)

VoIP technology overcomes this by converting voice-data into short, same-length packets each containing a few milliseconds of speech. The short packet-length facilitates buffering and easy re-sequencing, so that playback is to all intents and purposes immediate. Dropped packets may be skipped rather than being re-sent, with the VoIP codec smoothing over gaps in the packet sequence.

Again, RTP - Real-Time Transport Protocol - is used to synchronise the data stream between the two ends of the session using techniques developed for streaming media types, ensuring that data packets are decoded in the right sequence and that audio quality is maintained.

It is no co-incidence that the co-author of RTP is also the person behind the SIP protocol.

## VoIP Session Management Protocols

The following introduces VoIP session management protocols. These are the protocols that enable point-to-point services to be established and maintained, and provide for a number of other standard telephony requirements.

VoIP services need to emulate a number of conventions that have become embedded in normal telephony practises, from the basic, such a ‘making a call’, through to more elaborate services such as call transferring, putting calls on hold, call-forwarding, tele-conferencing, messaging, IVR systems, etc.

Currently the two main protocol suites used for VoIP are SIP and H.323. This article will focus on SIP however H.323 will be mentioned for completion.

### SIP

SIP stands for **Session Initiation Protocol**. It is an **IETF** standard that was developed by the co-author of RTP [Realtime Transport Protocol]. Originally developed for multi-point conferencing, it was quickly realised that SIP would also work for point-to-point conferencing - or phone calls. Created to describe the functionality required to start, maintain and terminate peer-to-peer communication sessions, SIP is designed to be interoperable with many other kinds of protocol and service to provide additional capabilities – in other words, it is an extensible architecture.

SIP provides for transmission of data types other than voice, including video and audio formats, and indeed was developed to accommodate multimedia capabilities.

Services have been devised using SIP including local and long-distance telephony, multimedia conferencing, virtual phone numbers, voicemail notification by email, and others too numerous to mention.

In keeping with its Internet-based heritage, SIP has much in common with HTTP and SMTP and indeed uses many of the same codes – for example, a SIP address is called a URI, meaning Uniform Resource Indicator, and is formatted similarly to an email address, i.e. <sip:user@sip.telcom.com>. SIP also has implicit compatibility with the DNS system that is the backbone of the Internet, enabling SIP-based systems to leverage of a great deal of existing infrastructure.

In common with HTTP and SMTP, SIP is written in plain text, providing for ease of use and analysis.

**Key SIP Terminology**

SIP operates in conjunction with other protocols such as:

- RSVP, which provides Quality of Service by negotiating optimal network resources on behalf of the voice bitstream
- RTP, or Realtime Transport Protocol, to ensure isochronicity and voice stream integrity, as mentioned previously

Other functionality can be invoked by a MIME-type capability which matches applications with filetypes to handle multimedia sessions. So a datastream carrying video information will be automatically handled by the appropriate video software on the users machine.

**Key SIP infrastructure:**

- SIP User Agents (UA) – are the end-user devices, whether these be VoIP phone handsets, 'softphone' applications on a PC, or PDA – whatever is used to create and conduct a SIP session
- SIP Registrar Servers - compile and track individual User Agents within their domain [e.g. sip:user1@theirdomain.com]
- SIP Proxy Server - pools SIP session requests, queries SIP registrar servers to locate the appropriate User Agent and forwards session invitations [=phone calls] to UAs in its domain or to another Domain Proxy Server as required
- SIP Redirect Server – may reside on the same physical machine as the above, and allows SIP proxy servers to direct SIP session invitations to external domains.
- SIP Gateway Server – interfaces between VoIP and the good old PSTN network so that VoIP calls can be connected to ordinary telephones

**Key SIP Commands**

- 'Invite' – an 'invitation' is the SIP equivalent of 'making a call', and is an 'invitation to communicate'. An invitation is invoked by the 'call' function of a VoIP phone and basically means 'do you want to talk?'
- 'Register' – the registration of a UA informs SIP registrar servers of the VoIP device whereabouts and availability

A complete list of SIP commands, notable for its brevity, is as follows[16]:

| SIP Method | Description |
|---|---|
| INVITE | Invites a user to a call |
| ACK | Used to facilitate reliable message exchange for INVITEs |
| BYE | Terminates a connection between users or declines a call |
| CANCEL | Terminates a request, or search, for a user |
| OPTIONS | Solicits information about a server's capabilities |
| REGISTER | Registers a user's current location |
| INFO | Used for mid-session signaling |

### Making a Call using SIP

When you 'dial' to make a VoIP call, an **invitation** is issued to the requested VoIP User Agent. The sending UA can either enter a URI [e.g. sip:user1@domain.com] or a plain old phone number, to send the invitation. The invitation is handled by the SIP proxy server which returns a '100' message to the calling UA, indicating that the call sequence has been initiated, while it looks for the receiving UA.

If successful, the invitation is answered by a '200' response from the remote User Agent which verifies the receiver is ready and that parameters are established for a communication session. At this point the SIP proxy hands off the session to the two endpoints which establish a Realtime Transport Protocol stream to carry the conversation.

### H.323

An alternative VoIP architecture, H.323 is a standard defined by the ITU, thus originating within the telecommunications industry rather than the Internet industry. Originally designed to accommodate multimedia, multi-point conferencing, H.323 is intended to provide a complete description of many sophisticated multimedia functions and is used in many high-end telecommunications switching devices.

SIP and H.323 compete in some respects, and are interoperable in others. SIP is gaining ground particularly among companies originating from the computer and internet industries, while H.323 is the more mature of the two and is well-represented in the telecommunications sector.

SIP enjoys some distinct advantages in terms of simplicity, as opposed to the far more complex H.323 standard, and in fact there is a 'format war' going on between these two with advocates from either side involved in sometimes acrimonious debate. Overall, the minimalist approach of SIP attempts to build only what is missing from existing communications technologies while synergising with other elements of the IP suite, whereas H.323 attempts, perhaps more ambitiously, to provide a detailed description of every facet of IP-based communications. One could almost characterise the two approaches as a protocol versus an architecture.

[16] Source http://www.sipcenter.com/sip.nsf/html/SIP+Signaling

### The VoIP Gateway

Basic VoIP can work as a peer-to-peer application between IP endpoints – meaning you and your friends can download a softphone and communicate with no intermediaries being required. But additional facilities are needed to make VoIP a real alternative in telecommunications environment still dominated by the PSTN systems, and also to integrate VoIP equipment with the PSTN world.

To interface with PSTN networks and existing equipment, a device called a **media gateway** (often called simply a 'gateway') is required. The gateway has a data network interface on one side, and a PSTN interface on the other, and its role is to intermediate between the two network types.

Gateways vary in size and scope. Industrial-strength, exchange-level gateways are used to intermediate large-scale networks, while single-port gateways, called 'Terminal Adaptors' or ATAs, are used to connect a single device or single PSTN services through an ADSL or cable modem. It is these household-level ATA devices that are set to become very popular in the next 12-18 months as broadband-equipped households and offices begin to use VoIP in earnest.

# Chapter 7: Introduction to VPN

Imagine you worked in a company with three offices, each 100 kilometres apart. Next, imagine that you needed to set up a communications link between all three locations so that confidential company data could be sent securely between them with no risk of interception. So you set about acquiring all of the cabling and other equipment necessary to build your own data links between your offices.

Once finished, such a set-up could be called a Private Network. But, as you can imagine, this would be a very expensive thing to create, install and maintain: several hundred kilometres of wiring, along with poles to string it on, or tunnels to bury it in, and the various kinds of switches and gadgets you would need to make it all connect.

Alternatively, you could hire your own dedicated lines from a telecommunications company. This would be considerably less expensive, because you would be able to lease or rent lines that were already installed. Nevertheless, research would show that this is still quite an expensive solution: leased lines do not come cheap.

Nowadays there is another alternative, and it is one that is becoming a standard all over the world. This method uses the ordinary telecommunications network to send and receive confidential information between each branch of the business, but it 'encodes' or 'encrypts' the information in such a way that no-one outside the business can intercept, interpret or read the data, even while it is travelling through the public network.

And this is the Virtual Private Network. It is 'virtual' because it uses public infrastructure, so it is not *physically* a private network. It sends all the data over the same phone network and satellite links that everyone uses. But because the data protection is so secure, it is, to all intents and purposes, private. Hence, VPN.

## How Hackers Operate

Hackers can do a lot of things to capture your confidential information. They can attach a computer to an unprotected point on a network and 'sniff network packets' using software tools, including commonly-available network utilities, to read and capture data.

A network password might be stolen, allowing an intruder access to your entire company network. Situations have frequently been discovered in which supposedly confidential company systems had long been compromised by users with stolen passwords.

Other kinds of threat include Denial of Service attacks, which are aimed not at penetrating the network but at bringing it down, by flooding it with so many requests that it is unable to respond.

Or hackers can analyse the network information contained in data packets to capture the address details of the originating and receiving computers, allowing them to imitate or 'spoof' a legitimate machine address, thereby fooling your network into thinking they are a part of it.

Even some methods of scrambling or 'hashing' your data may not provide complete security, as they might allow a hacker to get information about your network by determining the send-and-receive info, even if the content of the data packet is unreadable.

'Replay Attacks' occur when hackers capture a stream of data and then use it at a later time - 'replay' it – to gain unauthorised access, which is one of the reasons that 'time-stamping' access codes has become important. 'Man-in-the-middle' attacks are similar to the 'spoofing' attack and consist of an unauthorised user inserting him/herself into a communications sequence to become a trusted part of a supposedly secure communications session.

## The VPN Tunnel

These kinds of activities have created the need for VPN solutions. Current VPN technology allows two modes of secure data transmission, namely **transport mode** and **tunnel mode.** In transport mode the data packet's payload [=information content] is encrypted, with the IP [=Internet Protocol] header information being left unaltered. In Tunnel Mode, the entire IP packet is encapsulated, meaning that routing information in the packet header is also encrypted. Typically, tunnel mode is used for VPN gateway-to-gateway communication, while transport mode is used for host-to-host communication, a 'host' being defined as a sender or receiver of the information, and a 'gateway' as a device that monitors and manages network traffic.

Put another way, in a VPN tunnel-mode connection, remote computers are connected through the internet by creating a **tunnel** which allows data which is **encrypted** and **encapsulated** to pass between them. The 'encryption' applies to the data content, so that it will be scrambled ('hashed'); and the 'encapsulation' makes it impossible to determine the originating or receiving addresses of the data packets by disguising the 'from' and 'to' information contained in the packet headers.

As a result, if you are 'outside the tunnel', there is no way to understand the data being sent, nor can you determine where it is coming from, nor where it is going; it's all hashed to the point of being random zeros and ones.

Communications sent through VPN also have attributes such as **user authentication** and **data integrity** which confirm the sender ID and that the message content has not been altered in transit.

VPN connections are made and validated by means of a rigorous authentication process. It is important to grasp that an ordinary user-name and password log-on is not sufficient to establish a VPN tunnel; IPSec, explained below, is based on Digital Certification and is often maintained with the use of dynamic pass-codes which are changed at frequent intervals. Elsewhere we'll look at the specifics of how this is achieved.

### Common VPN Configurations

Common flavours of VPN are **site-to-site** and **site-to-end-user** configurations.

Site-to-site refers to VPN tunnels between building locations. In site-to-site configurations, the IT Administrator may establish the tunnel by installing VPN Gateways at each building.

Site-to-end-user configurations are generally installed to allow remote or off-site employees or trusted contacts to establish VPN communications with office systems.

## *Introduction to VPN Protocols*

In order to understand how VPN is implemented, you need to know something about VPN protocols[17]. It is important to understand that VPN consists of a number of methods, technologies, agreed practises and protocols in various stages of evolution; mature in some respects, and still developing in others.

[17] 'Protocol' - A standard procedure for regulating data transmission between computers

Oddly enough, when TCP/IP, the language of the Internet, was first invented, it never occurred to anyone that it might need to be secure; back then, in the 1970s, 'the enemy' was behind the Iron Curtain and 'hackers' were characters in horror movies. Internet commerce and domestic email were decades away and the main concern was to build a technology that could withstand military attack. Besides, no-one outside the Defence Department had any way of reading the data formats being developed for the Internet, and so security was a non-issue.

With the massive commercialisation of the Internet and the frightening prevalence of viruses, hackers, 'malware' and fraud, the need for security has since become abundantly clear.

**PPTP, L2TP and SOCKS**

Before looking at the main Internet Security Protocol, IPSec, we'll look briefly at PPTP, L2TP and SOCKS Firewalls, which are protocols and technologies supported by various vendors to provide secure communications.

PPTP [Point-to-Point Tunnelling Protocol] is a commonly-encountered, password-based VPN protocol. PPTP became available while IPSec was still under development and does not rely on, nor accommodate, digital certificate-based log-in authentication. Microsoft provides PPTP-based remote access clients with Windows 98/ME/NT 4.0/2000/XP and Win Server 2003 and recommends it as a secure tunnelling protocol provided it is used with suitable password practises.

L2TP stands for 'Layer 2 Tunnelling Protocol' and is a derived from Microsoft's PPTP and Cisco's Layer 2 Tunnelling Protocol. It does not in itself provide encryption and so is often used in conjunction with IPSec, hence 'L2TP/IPSec'. VPN solutions based on L2TP/IPSec have been implemented by many vendors including Microsoft, Cisco, Northern Telecom and Intel and are fully supported by NetComm.

Both PPTP and L2TP/IPSec are supported in Mac OS X V10.2 and above.

SOCKS is a secure communications technology based around a highly-configurable proxy server technology. SOCKSv5 is an SSL [Secure-Socket Layer] VPN, that is, it works on the Applications Layer rather than the Network Layer of the OSI Network Stack. It is said to be more suitable for individual remote access than LAN-to-LAN VPN connection. In practise, while SSL-level VPNs are secure, they are not nearly as 'transparent' to the end user as are IPSec-based tunnels.

**The IPSec Suite**

The growing problem of fraud, theft and hacking lead the Internet Engineering Task Force [IETF] to set about developing the IPSec ['Internet Protocol Security'] suite of protocols to enable secure internet connection. IPSec provides a framework for authentication, encryption, and encapsulation of sensitive data based on an end-to-end security model comprising a trusted link between two remote computers. Each computer handles security at its respective end, on the understanding that the medium over which the communication takes place may not be secure.

An overview of how IPSec protocols work will provide a good grounding for understanding VPN basics. Be aware that the mathematics of Data Encryption and Key Exchange are not easily explained or understood. It may not be rocket science, but it is computer science. You don't have to understand the maths behind encryption to form a concept of how it works. But you do need to get an idea of several interlocking parts of the solution, as, for example, key exchange depends on encryption and also on digital certification.

## Internet Security Association and Key Management Protocol - (ISAKMP)

The framework for authenticating communication between peers is described in the Internet Security Association and Key Management Protocol. Note that the concept of Key Management is intrinsic to the IPSec protocol suite. Key Management is effected via IKE (Internet Key Exchange) and PKI (Public Key Infrastructure), about which more later.

### Negotiating a Security Association (SA)

The first part of this framework, and the first step in creating a secure VPN connection, is in establishing a **Security Association**. According to Microsoft, 'A security association (SA) is the combination of a negotiated key, security protocol, and security parameters index (SPI), which together define the security used to protect the communication from sender to receiver.'

Thus **a Security Association is dynamic**: it is a session-based protocol that is created on the basis of variables which frequently change. This is part of the strategic architecture of IPSec, as static communications links are too easy to penetrate. The dynamism of IPSec ensures that even if a VPN tunnel is penetrated, the parameters will inevitably change, essentially 'sealing the tunnel' again.

A computer that is acting as a VPN server is often required to negotiate and maintain a number of simultaneous Security Associations. It is also possible to configure your Win 2000/XP/MacOSX computer with multiple Security Association profiles.

The establishment of a Security Association is managed by the formal process of Internet Key Exchange [IKE]. IKE ensures confidentiality and authentication through a sequence of steps which culminate in the creation of the secure VPN Tunnel according to agreed parameters.

In practise, system administrators must agree on and enter certain common factors [for example, encryption type, hash algorithms, network and machine address details, etc], prior to initiating the Security Association sequence. The SA process then follows this sequence:

### SA Phase I

First phase of an SA negotiation [Phase 1 or Main Mode SA] establishes:

the encryption algorithm[18] [DES, 3DES]
the hash algorithm [MD5 or SHA1]
user authentication [Kerberos V5, Certificate, or shared-key authentication]
The DH Group to be used for 'base-keying' material

If the two ends of the connection cannot agree on these terms, the SA negotiation will fail. If the two ends of the connection agree on these terms the negotiation proceeds to phase II.

The purpose of this phase is, then, to agree on the terms of the connection and also to provide security for the completion of the next phase.

### SA Phase II

Second phase of an SA negotiation [Phase II or Quick Mode SA] initially comprises policy negotiation, describing:

[18] Algorithm: A set of ordered steps for solving a problem, such as a mathematical formula or the instructions in a program.

The choice of IPSec protocol [AH or ESP = Transport or Tunnel Mode, see below]
the hash algorithm [MD5 or SHA1]
the [optional] encryption algorithm [i.e. 3DES or DES]
If this point is reached, an agreement has been negotiated and two SAs have been established: one for inbound, and one for outbound communication.

## *VPN Terminology*

Various terms in the SA process can now be discussed, particularly Encryption, Hash Values, and User Authentication.

### About Encryption

Encryption is the subject of the science of cryptography, and was originally practised by secret Babylonian religious sects anxious to evade the scrutiny of the authorities. Subsequently developed for military and diplomatic purposes, the word is derived from the Greek *kryptós* 'hidden' and *gráphein* 'writing'.

A very simple way of illustrating the encryption of a message would be that the Sender and the Receiver both agree to swap every letter in the alphabet for the letter that follows it. This is known as a **substitution cipher**. In this scheme, 'cat' becomes 'dbu' and dog becomes 'eph'. Of course, such a childish device would be completely transparent to an experienced code-breaker, but this kind of technique was where encryption started.

In modern cryptography, the bits of data representing the text of a message are encrypted by a program that transforms them by a factor such as a 168-character [=$2^{168}$] text string. In order to **decrypt** this text, the exact 168-length text string must be known.

A 168-bit key creates:

374,144,419,156,711,000,000,000,000,000,000,000,000,000,000,000,000

possible combinations that an intruder would need to guess in order to decipher your text![19]

This would be scarcely possible even for an expert code-breaker armed with the world's fastest supercomputer, particularly because the 'key' can be changed periodically, so that by the time it is calculated, it will no longer be valid. You might have a window of five minutes in which to guess this number. And while it is tempting fate to assert that no system can ever be broken, you could win the lottery a lot of times before breaking a code this complex!

Even keys of lesser numerical value are effective, as any encryption takes expensive computer-time and expertise to decrypt, and depending on the value of the encrypted information, is often not worth the effort of cracking – in other words, it is *computationally infeasible* to decipher it.

#### Encryption Methods: DES and 3DES[20]

In practise, encryption is generally applied via DES or 3DES (pronounced Triple DES). DES stands for 'Data Encryption Standard' and is a secret-key cryptography method that uses a 56-bit key. Based on a method originally developed by IBM and further refined by the U.S. National Security Agency, DES uses the **block cipher** method which breaks the text into 64-bit blocks before

[19] 'Understanding Virtual Private Networking' [.pdf] by ADTRAN Inc, p.6.

[20] Adapted from http://www.techweb.com/encyclopedia/defineterm?term=DES&x=24&y=16

encrypting them.[21] It is an efficient method which has become standard throughout the IT industry although it is now known that DES is not unbreakable, given enough computer power. For this reason 3DES is frequently specified.

If DES is used with a **secret key**, the key may be kept confidential and used repeatedly. Alternatively, a random key can be generated for each session, and transmitted to the recipient using a public key cryptography method such as RSA.

Triple DES increases security by extending the key to 112 or 168 bits, but requires multiple passes and takes more time.

Both formats are supported by NetComm. See also 'Key Encryption', below.

### One-Way Hash Functions

'One-way Hash Functions' are used to verify that a file has not been altered since being sent. This is achieved by calculating the total 'bit-value' of the original file and multiplying it by an arbitrary hash algorithm to create a **digest**. When the recipient receives the file, and performs the same function using the same hash algorithm, the **digest** will be the same unless the file content has been altered. Hashing the same message with the same algorithm will always produce the same digest. Change the message by one character, and you'll get a completely different digest. Note also that the length of the digest is always the same, regardless of the length of the original input. This is one of the reasons why it's called 'digest'.

Examples of hash algorithms are MD5 and SHA-1. Both of these formats are supported by NetComm.

The point about 'one-way functions' is that they are easy to calculate going forward – that is if you have the input values – but very hard to calculate by going backward – that is, to deduce the input from analysis of the output.

## User Authentication Methods I: PAP, CHAP and MS-CHAP

Having encrypted your data for security, and hashed it for authentication, the other requirement for establishment of a VPN is two-way authentication: authentication of the user ID to the service, and authentication of the ID of the service to the user.

The simplest, and earliest, forms of User Authentication were provided by basic User Name and Password functions. These were originally developed as part of the Point-to-Point Protocol [=PPP], with the most basic format being Password Authentication Protocol [=PAP]. This method, however, transports the User ID and password in clear text and is thus too easy to intercept. Improvements included Challenge Handshake Authentication Protocol [=CHAP] and the several Microsoft enhancements to this, including MS-CHAP and MS-CHAP v2.

### MS-CHAP v2

From Microsoft's 'VPN Overview' White Paper: 'MS-CHAP v2 is an updated encrypted authentication mechanism that provides stronger security for the exchange of user name and password credentials and determination of encryption keys. With MS-CHAP v2, the Network Access Server [=NAS] sends a challenge to the Access Client that consists of a session identifier and an arbitrary challenge string. The remote access client sends a response that contains the user name, an arbitrary peer challenge string, and an encrypted form of the received challenge string, the peer challenge string, the session identifier, and the user's password. The NAS checks the response from the client and sends back a

[21] This technology was a 'prohibited export' in the USA for many years due to classification as 'munitions'

response containing an indication of the success or failure of the connection attempt and an authenticated response based on the sent challenge string, the peer challenge string, the encrypted response of the client, and the user's password. The remote access client verifies the authentication response and, if correct, uses the connection. If the authentication response is not correct, the remote access client terminates the connection.

'Using this process, MS-CHAP v2 provides mutual authentication: the NAS verifies that the access client has knowledge of the user's password and the access client verifies that the NAS has knowledge of the user's password. MS-CHAP v2 also determines two encryption keys, one for data sent and one for data received.'

## User Authentication Methods II - EAP

User-authentication protocols requiring more than a simple username and password are summarised under the heading of **Extensible Authentication Protocol** [=EAP]. According to Cisco Systems, 'EAP provides a standard mechanism for supporting various authentication methods over wired and wireless networks.' EAP supports authentication types such as Generic Token Card, One Time Password (OTP), MD5-Challenge, Transport Level Security (TLS) for smart card and certificate support, as well as future technologies not yet devised or in use.

### Two-Factor Authentication

Two-factor EAP authentication systems require users to enter more than a username and password and are are similar to an ATM card and PIN number – 'something you have and something you know'. You **have** the card, and you **know** the PIN – and access can only be gained with both.

Two-factor authentication may be provided by way of dynamic pass codes which are distributed in addition to the normal User ID and Password log-in. Sometimes these are provided via credit-card size devices which display a dynamic pass code by LCD display; in other cases the information is provided via SMS or mobile phone call. Some systems require the presence of a PC card or USB key in addition to the user name and password. In all cases the user must present something they have, and something they know, to access the system.

For example, Vasco Systems' 'DigiPass' solutions are credit-card-sized devices which provide a dynamic passcode that must be entered to access a secure system. But in order to retrieve the passcode, the user must know the password for the DigiPass. This means you must have the device AND know that password to retrieve the passcode. If you have either the device OR the password by themselves, you will not be able to gain access.

## PKI – Public Key Infrastructure[22]

In practise, the principle of Key Cryptography is managed through a set of protocols called **PKI**, or **Public Key Infrastructure**, which is based on **digital certificates** and **Certification Authorities** (CAs). This is necessary because for such a scheme to work it must be based on a foundation of trust and transparency while providing common methods and definitions to enable diverse users and systems to interact. PKI provides a set of definitions, practises and protocols that are available for commercial and individual use through the use of digital certification and are implemented through a system called Internet Key Exchange [IKE].

[22] http://www.microsoft.com/technet/prodtechnol/windows2000serv/deploy/confeat/cryptpki.mspx

You might be tempted to ask what these 'keys' look like. Actually, they are simply sequences of numbers and letters. They only work by virtue of being almost impossible to guess. Hence their value is derived from the mathematical and regulatory framework in which they exist.

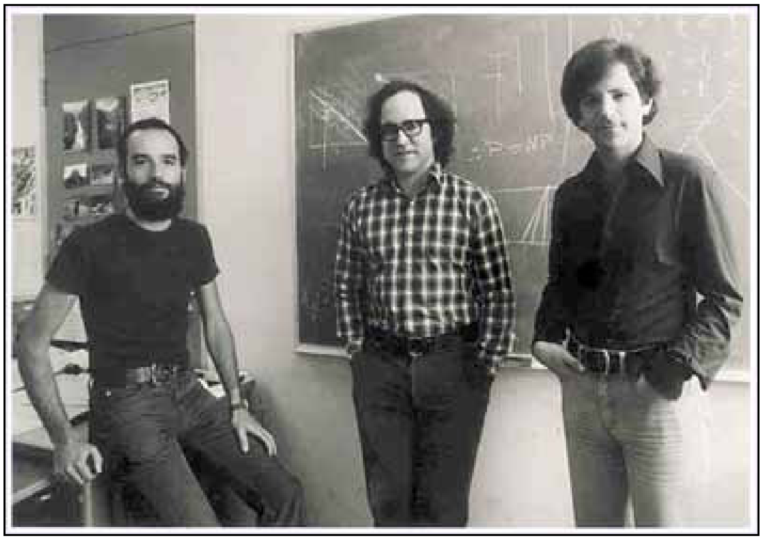
### Shared Key Encryption: One Key

Shared key encryption [also called symmetric key, secret key and session key] uses a single key common to both the sender and receiver and to encryption and decryption. An obvious shortcoming of this method is that if a third party knows the key, it is no longer effective. If you wish to send encrypted text to someone using Shared Key Encryption, they must have the Shared Key in order to be able to decrypt the message. This means you can't send the Shared Key in plaintext or by other insecure means without compromising its security.

### Public-Key Encryption: Two Keys

Public-key encryption [also called asymmetric-key encryption] uses two keys: a public key and a private key, which are mathematically related. In public-key encryption, the public key, as the name implies, can be shared or published, so long as the paired private key remains private. Data encrypted with the public key can be decrypted only using the private key. Data encrypted with the private key can be decrypted only using the public key.

While effective, their use of very large numbers makes Public-Key Encryption relatively slow. However Public Key algorithms are often combined with other cryptographic elements, for example, with hash algorithms to generate digital signatures, or combined with Shared Keys for key exchange.



Common types of Public Key Algorithms include:

**RSA** – for Rivest-Shamir-Adleman, the inventors of the algorithm (pictured). Supports both Public Key and Digital Signature functions. Widely used and supported by most major vendors.

**DSA –** for digital signatures; incorporated by the US National Institute of Standards and Technology.

**Diffie-Hellman**[23] – for key exchange only. The 'DH Group' stage in SA Phase I refers to the Diffie-Hellman exchange sequence. Basically the DH algorithm allows

[23] Named after Whitfield Diffie and Martin Hellman, mathematicians at Stanford University. See http://livinginternet.com/i/is_crypt_pkc_inv.htm

a secret key to be generated between two computers on the basis of exchange of numbers which in themselves are not secret. This is called a 'shared secret' technique.

In practise, IKE settings are generally dealt with by IT administrators when configuring VPN gateways, servers and IP Security Policies either for users or for LAN-to-LAN gateway installations. This will be illustrated below using NetComm gateway setup as an example.

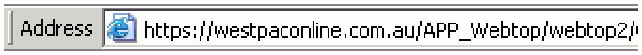
**RADIUS Servers**

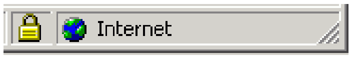
On the server end, user authentication is being standardised around the Remote Authentication Dial-in User Services [=RADIUS] server suite, [24] a protocol-based suite which covers user authentication, authorization and accounting and is frequently configured as the central authorisation hub of a VPN.

**Digital Certification Systems and Secure Socket Layer Communications**

At the root of the PKI system are the Certification Authorities, or CAs.[25] These organisations are licensed to produce **public key certificates.** The important thing to grasp about digital certification is that it is based on a **chain of trust**: at the beginning of the chain, there is an identity that has been verified by indisputable means, which is conferred through each link in the chain to create the integrity of the Digital Certificate.

When a browser session is secured by a Digital Certificate, the URL will commence with <https://> and the browser window will display the 'padlock' icon that denotes an SSL [Secure Socket Layer] transmission. Examples are as follows:


Address https://westpaconline.com.au/APP_Webtop/webtop2/



When these indications are made in the browser window, users are assured that confidential information they enter such as their name and credit-card details will be encrypted for secure transmission.

According to VeriSign, a leading Certificate Authority, 'Digital IDs are the electronic counterparts to driver's licenses, passports, and membership cards'.[26]

A Digital Certificate attached to a secure website verifies that the site really is what it purports to be. The need for this is illustrated by the current spate of 'phishing' scams whereby hackers send bogus emails, purportedly from a bank, luring users to a counterfeit website which is used to capture their confidential log-in information.

Browser programs will display an alert if the digital certificate name doesn't match the name on the secure site. One would hope this would be sufficient to prevent users from revealing their banking details to a bogus site….

So, typical Digital Certification services include:

[24] RADIUS is described in RFC 2865, "Remote Authentication Dial-in User Service (RADIUS)," (IETF Draft Standard) and RFC 2866, "RADIUS Accounting" (Informational).

[25] A complete list is at http://www.pki-page.info/

[26] https://digitalid.verisign.com/client/help/id_intro.htm

### Server ID

This enables web servers to operate in secure mode. The Server ID identifies the server as well as providing encryption for data passed between client and server during the secure session.

### Developer ID

This provides certification for software developers, and is often used in conjunction with Microsoft's Authenticode™ software validation technology to verify the source of the download.

### Personal Digital ID

This is used by individuals to validate and authorise messages and may be required to access certain secure systems.

The following functions are supported by Digital Certification:

### Digital Signature

Digital signatures enable authentication of digital messages and files, assuring both the identity of the sender and the integrity of the message itself to the recipient.

A VeriSign example: Suppose Alice wants to send a signed message to Bob. She creates a message digest by using a hash function on the message. The message digest serves as a "digital fingerprint" of the message; if any part of the message is modified, the hash function returns a different result. Alice then encrypts the message digest with her private key. This encrypted message digest is the digital signature for the message.

Alice sends both the message and the digital signature to Bob. When Bob receives them, he decrypts the signature using Alice's public key, thus revealing the message digest. To verify the message, he then hashes the message with the same hash function Alice used and compares the result to the message digest he received from Alice. If they are exactly equal, Bob can be confident that the message did indeed come from Alice and has not changed since she signed it. If the message digests are not equal, the message either originated elsewhere or was altered after it was signed.

### Digital Time-Stamping Service

DTS is required to validate the exact creation and modification dates of a digital document. This can be used, for example, in authorship, copyright or patent law, or to monitor time-dependent access to secure systems.

### Certificate Revocation

Various grounds may give rise to the need for a Digital Certificate to be revoked. These include key compromise, issuing authority compromise, employee departure, business closure, and so on. Revocation information is published in a Certifica Revocation List which CAs are required to maintain as part of their function. If a browser encounters a Revoked Certificate it will display an alert to that effect.

### Kerberos Authentication

Microsoft has adopted Kerberos Authentication, originally developed by MIT, as a method for implementing user authentication. The three main aspects of the Microsoft implementation of Kerberos comprise the Key Distribution Centre [KDC], the client user and the server with the desired service to access. Kerberos authentication is based on the principle of the user being given a time-stamped

Ticket which provides access to specific network resources as a validated user, for a specific time period.[27]

### Operating System Support for IPSec

Microsoft has enabled IPSec Policy Implementation[28] as a high-level feature of Windows 2000 and XP, having jointly designed several aspects of the IPSec implementation with Cisco Systems. In practise this means that IPSec authentication, connection and encryption are implemented via the Microsoft Management Console (MMC) and the Local Security Policy Control Panel. This implementation allows the user or administrator to create a named Connection Profile and to select encryption and authentication protocols.

Support for IPSec, PPTP and L2TP is provided in Apple's OSX and is implemented in various versions of UNIX also.

### Security Policy

Once a Secure Connection is established, the degree of authentication required to access network resources has to be decided. This establishes what is known as the **Security Policy**. It is important to understand that while the Security Policy may rely on the framework of protocols offered by IPSec, the Policy in itself is a result of the degree of security deemed appropriate by the Company or by the Network Administrator. The policy determines the use of IPSec protocols.

The Administrator may vary the degree of authentication required for individual users to log on to the VPN, and also the frequency with which session keys are refreshed between the end-points of the Tunnel.

In a some environments, users may log on with a static user-name and password which may change rarely or at the user's discretion. Similarly, the VPN settings at the VPN gateway may remain unchanged for long periods of time. In other places, complex passwords may be a requirement, along with dynamic pass-codes.

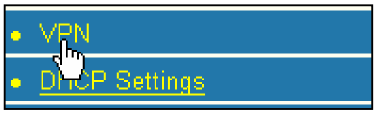
[27] Microsoft's intro to Kerberos is at http://www.microsoft.com/technet/prodtechnol/windows2000serv/maintain/security/kerberos.mspx There is a non-technician's overview in the form of a dialogue at http://web.mit.edu/kerberos/www/dialogue.html

[28] http://www.microsoft.com/technet/prodtechnol/windows2000serv/howto/grpolwt.mspx

Illustration: Basic Setup of point-to-point VPN

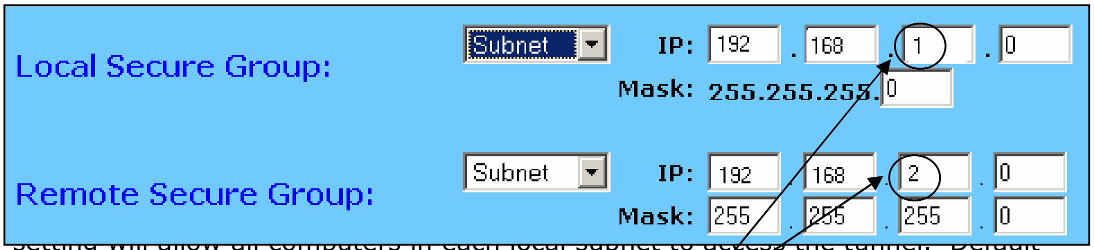
This summary illustrates how to set up a VPN between two locations 'A' and 'B' using either NetComm NB5540 or NB5580 VPN routers. This example only provides details for default settings; for details on other configurations, please refer to the manual.

**1.** Select VPN from Main Menu:



**2.** Click radio button to **Enable** a VPN tunnel and **name** it for ease of reference. The name in this field is only a reference and is **not** a VPN parameter. Name cannot contain spaces or special characters (^#@ etc).

**3.** Address Parameters for End 'A'



Local Secure Group: Subnet IP: 192 . 168 . 1 . 0 Mask: 255.255.255.0

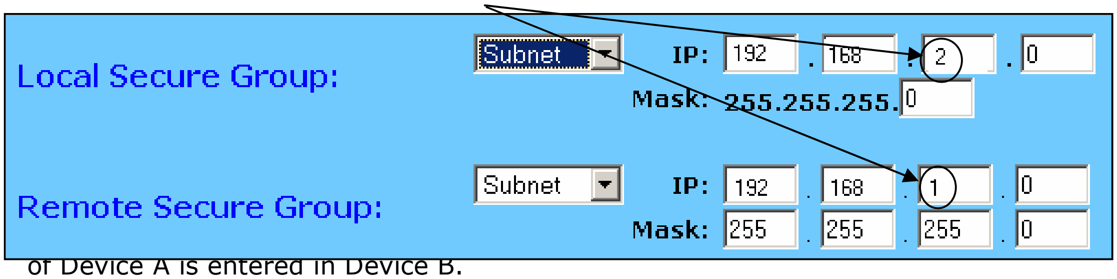
Remote Secure Group: Subnet IP: 192 . 168 . 2 . 0 Mask: 255 . 255 . 255 . 0

setting will allow all computers in each local subnet to access the tunnel. Default value of '0' is to remain in the last octet of IP and Mask fields.

**4.** Address Parameters for End 'B'

Address Parameters are reciprocal:

Local Setting for A = Remote Setting for B Remote Setting for A = Local Setting for B



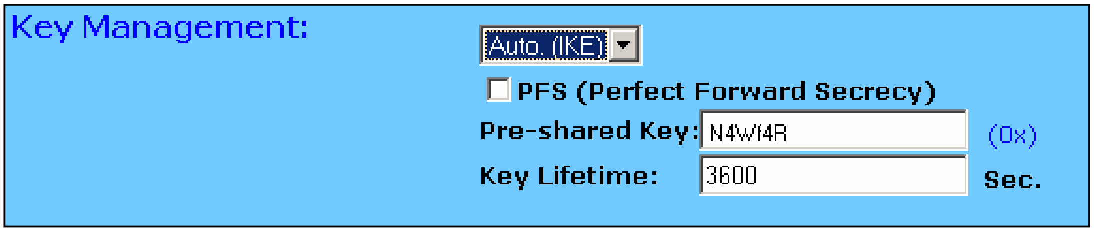
Local Secure Group: Subnet IP: 192 . 168 . 2 . 0 Mask: 255.255.255.0

Remote Secure Group: Subnet IP: 192 . 168 . 1 . 0 Mask: 255 . 255 . 255 . 0

of Device A is entered in Device B.

**6.** Encryption, Authentication

These **must** be the same values at each end. For security purposes, 'disable' is **not** recommended. 3DES is more secure than DES but may result in slower connection.



Key Management: Auto. (IKE)

□ PFS (Perfect Forward Secrecy)

Pre-shared Key: N4Wf4R (0x)

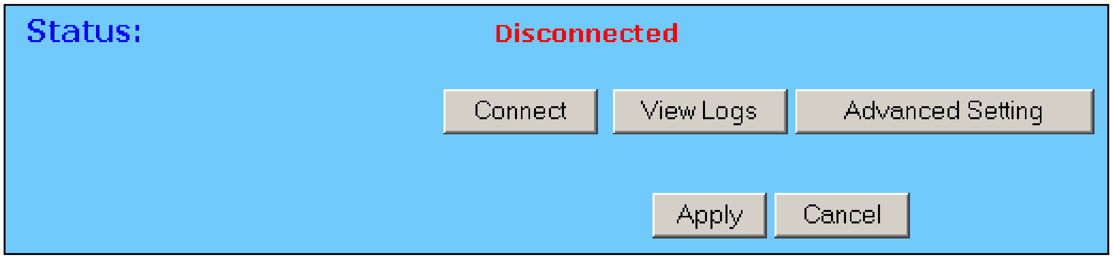
Key Lifetime: 3600 Sec.

**7.** Key Management

Key Management: Leave as default [Auto. IKE].

- In Pre-Shared Key, enter a sequence of alphanumeric characters. These must be same at each end of the VPN; no spaces and no special characters allowed. **DO NOT SEND KEY BY UNSECURED EMAIL**.
- Key Lifetime: leave as default.
- PFS: leave unchecked.

**8.** Connection



Status: Disconnected

When both ends of tunnel are configured:

Click **Apply** to save settings

Click **Connect** to establish VPN Tunnel

When connection successful, status indicator will change to ‘Connected’.

…[end]