case communications

The pocket book of

ETHERNET SWITCHING

Introduction

As companies and institutions have adopted computer technology, local area networks (LANs) have expanded in both size and importance. Networks have grown from being simple connections between a few machines to massive building-wide installations that connect to world-wide networks.

As network usage increases in size and complexity, network capacity must be increased otherwise problems such as slow response and lost data will begin to reduce the performance of the network.

Ethernet switches have been designed to alleviate the traffic problems that can arise on crowded LANs. Switches enable you to break up your network into many segments. Switches segments do not interfere with each other for local traffic, but each can send traffic easily to the others when needed. Ethernet switches are able to supply up to 10Mbps to each desktop without any upgrading to workstation network hardware or cabling.

This pocket book will help you to understand Ethernet switching. It will help you decide when a switch should be added to your network and which kind of switch is best for you. It starts with a history of Ethernet LANs, traces the development of switching and explains how it is different from other LAN networking equipment (such as bridges, hubs, and routers).

Specific issues surrounding Ethernet switching will also be discussed so that you can get a realistic idea of what switching will do for your network. Benefits and issues unique to switching, such as management, virtual LANs, filtering and security will also be explained.

A brief history of LANs

Local Area Networks (LANs) were first implemented in the late 70s as a way of interconnecting PCs. Data and documents could be moved between users without requiring paper thus coining the phrase “the paperless office”. At this time the technology used was only intended to connect devices within a single building, hence the use of the term “local”. Networks being implemented now support not only traditional LAN traffic, but also voice and video, giving rise to the ear of true multimedia communications.

To connect LANs together, Wide Area Networks (WANs) were used, which were based on radically different technology. Over the last few years the difference between LAN and WAN technologies has blurred.

How Ethernet works

The most common LAN technology is Ethernet. Originally, Ethernet was designed to enable many Personal Computers (PCs) and other devices to communicate efficiently over a single wire. To minimise the cost of ownership Ethernet was created to be inexpensive to install and easy to maintain.

How Ethernet works

To understand how Ethernet works, consider an Ethernet segment to be a large concrete pipe with lots of little pipes connected to it. Each little pipe represents a station or user. When a station wishes to send data, it listens to hear if there is any traffic on the big pipe. If there is none, the sender shouts the address it wants to speak to and then starts sending a message. As it is sending, the sender listens to the echo in the pipe. If the echo is clear, then the sender knows the signal is OK and continues shouting into the pipe. If the signal has been corrupted (for example, by someone else starting to shout down the pipe at the same time)) then the sender will hear an incorrect echo. In this case, the sender stops transmitting for a random time interval and then tries again to send the entire message.

It takes the data (talk) some time to move along the segment (pipe) before all users/stations connected to the segment become aware that the main pipe is busy.

Ethernet is defined as “busy” once 64 bytes of data have been sent. All other stations then become aware that the segment is busy and will hold off sending until the transmission is complete.

Ethernet addressing

The original inventors of Ethernet realised that, to be successful, it would have to be simple to implement and inexpensive to run. It needed to have a simple system for addressing (unlike WANs at the time, which were causing significant problems with their complex addressing schemes). Therefore Ethernet is based on simple Media Access Control (MAC) addressing, which assigns a unique MAC address to each device connected to the network.

Ethernet development

At first, Ethernet LAN installations normally consisted of a single section of cable (usually a 10-Base-5 cable, which typically is a yellow co-axial cable approximately 1 cm in diameter). This cable passed through each department in a building. PCs were connected to this wire by “Taps”, which enable PCs (or “stations” as they are called) to be connected to the “Ethernet highway”. At this time, most of the PCs were working as simple terminal emulators (just sending commands to a program running on a corporate host or a mainframe computer). Very little local processing was done locally, which meant that network transmissions were small and the 10Mbps of Ethernet capacity was more that adequate.

As the PC became a higher performance platform, cracks started to appear in the single segment Ethernet technology. PC users started to use the power of the PC for creating memos and letters locally. Entire documents were then communicated via electronic mail to other stations on the "data highway". This in turn caused traffic on the "data highway" to slow down causing bandwidth requirements to increase, so users noticed that traffic on the "data highway" was slowing down, just like motorway traffic when the number of cars increases.

Ethernet had a similar problem. As the number of users increases, so does the change of a collision, which slows down data transfer. The answer was to break the single segment into multiple segments using bridges to control the flow of data.

Structured solutions for larger networks

Soon, nearly every employee in the building required a LAN connection. Network Managers realised that most of the costs associated with operating a LAN were in relocating users when their departments moved, as all the LAN cabling had to be relayed. IT was for this reason that structured cabling came about. Structured cabling enables a computer network to reach each user's desk like telephone and electricity wires do. It makes moves and changes very easy and, more importantly, less expensive.

Traditional LANs used a bus architecture, with all network stations connected in a string on a single length of expensive co-axial cable. Structure wiring uses a star configuration, with a dedicated piece of inexpensive cable extending from the user's station to a central hub/concentrator. The hub enables all of the stations connected to it to communicate as a single LAN segment.

The cabling used in structured cabling is called unshielded twisted pair (UTP) cable. UTP cable is exactly the same as telephone cables, so they are inexpensive and readily available.

The hubs on each floor can be connected to one another through a "backbone" which runs up a vertical shaft within the building. This creates a building-wide structured cabling system. The backbone normally uses fibre cabling.

This configuration makes it easier for a user to move his or her computer to a new desk and plug it straight into the network connection that is already there.

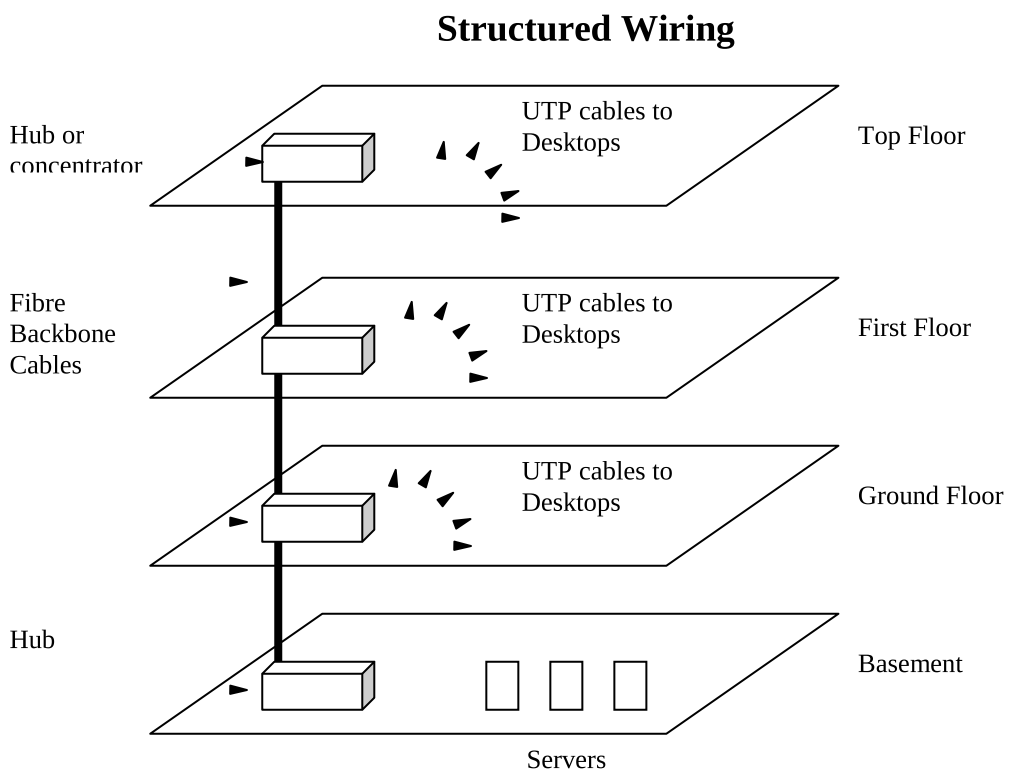


Figure 1: *Structured cabling showing fibre on the backbone and UTP to the desktop*

Enlarging LANs with hubs

AS mentioned above, the structure wiring approach requires that all stations communicate via a network component called a hub. All stations that are connected to the hub communicate as a single LAN segment. This makes it easy to provide network connections to a large number of people, even when they move frequently.

Unfortunately, that is all that traditional hubs provide – a way of connecting more users to the same segment, forcing all users to share the same bandwidth (this is like adding more on-ramps to a motorway which creates more traffic). Hubs do not normally include any intelligence. All signals that come in are immediately sent out of all ports to all stations.

Network Managers realised that this approach would eventually reduce network performance, but accepted that this problem would be solved later. They used hubs as a way to solve the immediate problem of providing network access to everyone who needed it.

The function of a hub is simple. For his reason managers purchased hubs that could provide the greatest number of connections for the lowest price. In other words, they purchased the hub that had the lowest price per port.

At first, chassis-based hubs were the only available devices, but then stackable hubs started to appear. Because of their very low price per port, stackable hubs have become very popular. They generally offer a very attractive price per port and for small networks provide similar levels of functionality to chassis-based solutions.

Segments

The concrete pipe example mentioned earlier demonstrates that Ethernet is a shared media.

Thought the speed of Ethernet is 10Mbps, the actual data throughput drops (due to overheads and collisions) after traffic increases past a certain point. As a general rule, maximum throughput on a shared segment is achieved when the network is around 40% utilisation, the maximum throughout available to each user is 200kbps. The lesser the number of users per segment, the greater the available bandwidth for each one.

To provide higher bandwidths networks must be divided into more segments. Each segment has 10 Mbps to be divided amongst the users connected to that segment, with optimal throughput provided at 40% utilisation.

If a segment has only one station, as it possible with switching, then no collisions will occur and utilisation can reach 100%. This equated to a full 10Mbps of throughput per user.

Segments using identical LAN protocols can be linked with either a filter bridge, switch or router. To connect dissimilar LAN segments (for example, Ethernet to Token Ring), a router must be used.

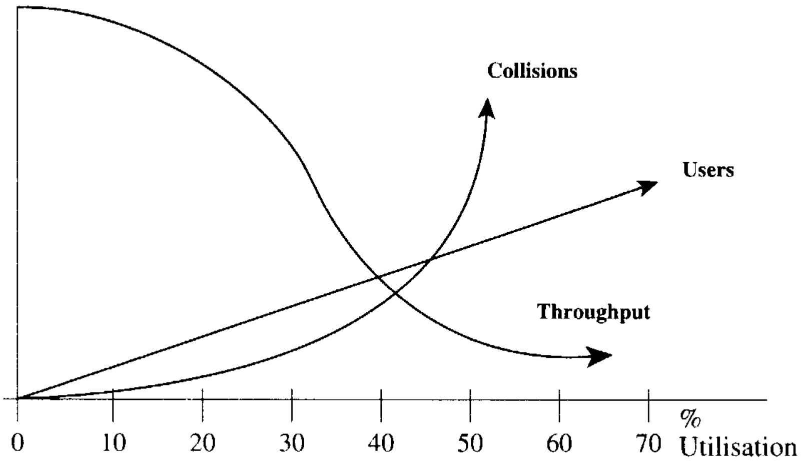


Figure 2: *Once a segment utilisation reaches 40% the throughput decreases substantially*

Bridges

Bridges are used to connect two or more LAN segments together. They often include a sorting function that enables the bridge to only send frames out to the segment that contains the destination station.

Bridges work at layer 2 of the OSI 7-layer communications model, making them protocol independent (For more information about the 7-layer model, refer to Case Communications *Pocket Book of OSI*). Bridges look at the MAC address to make forwarding decisions and do this without changing the Ethernet frame. Bridges are normally processor-based and can support approximately seven to eight segments.

When a bridge receives a frame, it reads the destination address. It then looks in an internal address table to decide which segment contains the station that should receive the frame. If the recipient is on the sender’s segment, then is does nothing. If the recipient is on another segment, then it sends the frame to that segment. This enables bridges to isolate traffic between segments.

When a bridge is first turned on, it does not know anything about then network. As it operates, the bridge builds its address table based on the addresses is has already seen. When a bridge receives a frame which is addressed to a device not listed in the address table, it will broadcast the frame (shout on all segments) and note which segment responds.

Some frames come into the bridge addresses as broadcasts/multicasts, which means that they are forwarded to all segments in the network.

If too many stations send broadcasts at the same time, then a “broadcast storm” results. This causes the entire network to become saturated with broadcasts, which all stations must read and process, this slows station performance. Unfortunately, bridges rarely do anything to stop broadcast storms from propagating over the entire network.

Routers

Routers perform all the functions of bridges and more. They are used to improve network segmentation, and to route between dissimilar LANs and also route to wide area connections.

Routers always use a store and forward technique, which means the entire frame is read into memory before it is sent on. Routers read all the layer 2 information and part of layer 3 (IP Address). They are able to identify the protocol in use from the LF Field. This means that routers work at level 3 of the 7-layer model, so (unlike bridges) they are protocol dependant.

A Router strips out the layer 2 and address headers and imbeds them into a new frame, which will be transmitted on a different type of LAN segment. This means that routers can be used to connect dissimilar LANs (for example, Ethernet to FDDI).

It takes time for the router to store the frame, read the layer 2 information, repack the data and send it on. This introduces delay which users normally perceive as slow response times and stops routers being used for multimedia.

Like a bridge, a router uses address tables to send the frame to its correct receiver. A router can also communicate with other routers in the network to learn the locations of different stations. This ensures the router sends the data to the correct destination along the shortest possible path. However it has an overhead on both the network and processing power within the router.

All routers "chit-chat" with each other and update their router tables in order to use the best path (route) between each segment. This means that routers must have a greater amount of intelligence than a bridge, and also require more bandwidth for the updates. Therefore, they require more processing power and generally sell at a higher price than bridges. However, because of the way they work, routers prevent broadcast storms from propagating through a network, thus providing improved network performance.

As organisations looked for higher performance at the core of their networks, vendors were arguing as to who had the highest performing router (measured, for example, by how many packets per second (PPS) it could support). At this point organisations were looking for performance on the backbone and they were generally prepared to pay for it.

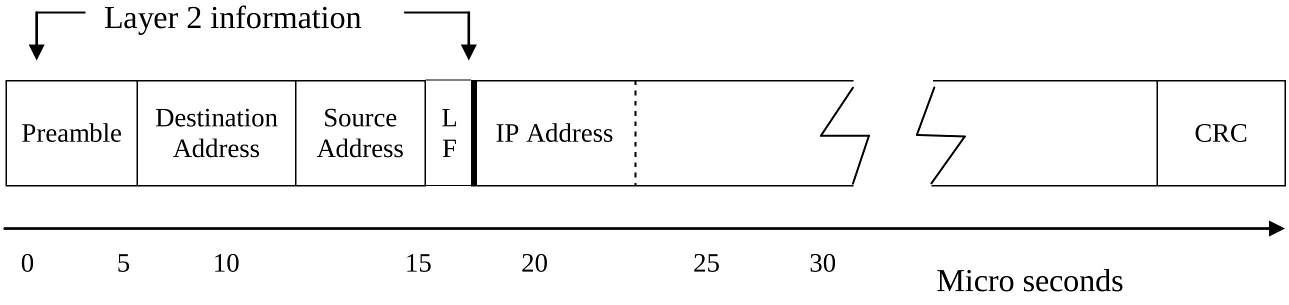


Figure 3: *An Ethernet frame is divided into several standard fields*

Switches

Ethernet switching was created to combine the fast throughput of bridges with some of the intelligence of routers. Like routers, they work by dividing the network into a number of segments, each of which can operate without interference from traffic local to any of the other segments.

Switching is done at level 2 of the 7-layer model – the same level as bridging.. AS it is done at level 2, the MAC address is used, which is independence of the protocol address. Like a bridge, a switch learns which addresses reside on each of its ports and then switches the data approximately. A switch can be designed using either a conventional microprocessor or dedicated ASIC technology (described later).

Switching provides higher network performance and introduces a number of new benefits. IT offers the possibility of single-user, collision-free segments that provide 10Mbps throughput to the desktop. Several types of switches are available for supporting both workgroup and backbone structures with uplinks to Fast Ethernet and FDDI.

Switches enable Network Managers to divide networks and virtual LANs. Virtual LANs make changes easier and can improve performance, prevent broadcast storms, and improve security. Switching also introduces some new challenges, such as how to manager and monitor for problems in a network divided into many independent segments.

	7	Application	
	6	Presentation	
	5	Session	
	4	Transport	
OSI/DECNet/IPX/IP Address	**3**	Network	Routers
MAC Address	**2**	Link	Switches Bridges
	1	Physical	

Figure 4: *Basic representation of the OSI 7-layer communications model. Each networking device operates using a specific layer to decide where frames should be sent. More complex devices operate on the higher layers.*

Switch Hardware

There are a number of ways an Ethernet switch can be designed. Since Ethernet switches (unlike hubs) are new, vendors are still experimenting with different types of hardware architectures. Variations exist in the forwarding techniques, switching technique, control hardware and buffering schemes.

Microprocessor-based architecture

A switch requires intelligence to manage switching functions. It needs to make and keep address tables, read addresses on incoming data frames and decide which frames must be forwarded to where. Many manufacturers use general purpose off-the-shelf microprocessors for this function.

A microprocessor is powerful and flexible, but expensive. So each switch generally contains switching, buffering, and management functions. This means that the switch can become overwhelmed and therefore slow down and lose data during periods of heavy traffic.

As a rule, a processor-based switch can support around eight to ten Ethernet ports. If more ports are required, then additional processors must be added to the switch. This makes processor-based switches non cost effective.

ASIC-based architecture

ASIC stands for Application Specific Integrated Circuit. This is a custom-designed silicon chip that performs a specific function (in this case, switching and/or other associated functions). Since each ASIC is optimised for a specific task, it is more efficient than a processor. Once designed, each ASIC is simpler, and therefore faster and less expensive, than a general-purpose processor.

ASICs can generally support a greater number of ports with less delay, as all the switching is done in the hardware. With a processor, all frames have to be examined and passed through Random Access Memory (RAM). An ASIC switch can normally support around 40 ports.

Several dedicated ASICs can be included in a switch design to manage different components of the switch's operation in parallel with one another. For example, one ASIC to manage address tables and switching functions, another for managing the buffer, and a third for monitoring traffic with RMON at each port. If a design requires more ports than additional port ASICs can be added for relatively little cost.

Buffers

Switches must use buffers to store data as it enters and/or leaves the switch. This enables the switch to transfer frames through its internal architecture and to connect ports that operate at different speeds. For example, 10Mbps workstations to a 100Mbps FDDI/Fast Ethernet backbone.

Frames can be buffered at the input port, at the output port, or a combination of the two. The most efficient method is to use output buffering. By keeping the input ports clear, incoming data can always be sent immediately to another output port, where it will be buffered only as long as required for the receiving station to read it.

Output buffering helps the switch to perform better under loaded conditions. Switches with output buffering have efficiency rates in excess of 98%, while input buffer based solutions are limited to 50-60% efficiency.

> Some switches are designed so that each port has a fixed amount of RAM to buffer frames (see **Figure 5**). This is called static buffering. Unfortunately, this results in wasted resources, as empty RAM is reserved at each port when it is not needed. It also means that one port can run out of buffer capacity while another has an excess of space. For example, a port connected to a server will have no more capacity than a port connected to a PC carrying out terminal emulation.

More efficient designs use a system that assigns buffer space dynamically to each port, drawing from a shared RAM pool (this is sometimes called an "elastic buffer"). Each port uses only the RAM it needs, when it needs it. By using a RAM pool, less overall RAM is required.

Static Memory

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

Example showing static memory allocation in a 16 port switch

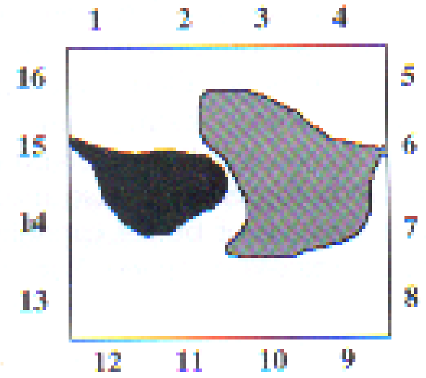


Figure 5: *Example showing how dynamic memory is utilised in a 16 port switch. Note greater flexibility and therefore lower memory usage*

Half/Full duplex

The difference between half and full duplex communication can be compared to the difference between speaking via an intercom or a telephone.

On an intercom (or ship radio), only one person at a time can speak while the other can only listen until the speaker is finished. Often, the speak must say “Over” to tell the other party that the line is free. This is half-duplex communication.

On a telephone, both parties can speak and listen to one another at the same time. This is full-duplex communication.

Until recently Ethernet had always worked in half-duplex mode as all stations had to communicate over a single cable, capable of resolving only one signal at a time.

More recently, LANs have been assembled using unshielded twisted pair (UTP) cables in structured cabling installations. UTP is also used for telephone installations. It was chosen for structured cabling in LANs because it is inexpensive and readily available.

Since UTP cables have been designed for telephones, they are capable of full-duplex communication. They contain two wires for receiving data, unlike co-axial cable (such as 10BASE5 and 10BASE2), which uses a single wire copper core.

In a traditional LAN environment UTP is used to connect workstations to a hub, which is connected to a backbone or a coax-type cabling system. Since the hub shares all data with all ports, the system is still limited to half-duplex.

However, in a switched environment, bandwidth is not shared, so full-duplex communications is possible over UTP cabling. This results in doubling the bandwidth from 10Mbps to 20Mbps.

Full-duplex is especially useful for connecting devices that transmit large amounts of data in both directions, such as in switch-to server or switch-to-switch connections. The increased bandwidth of full-duplex connections is easily capable of supporting interactive and two-way multimedia applications.

In order to run full-duplex, the PC interface card must be able to support full-duplex working. Therefore the optimum configuration without changing existing interface cards is to run servers full-duplex and leave clients running half-duplex.

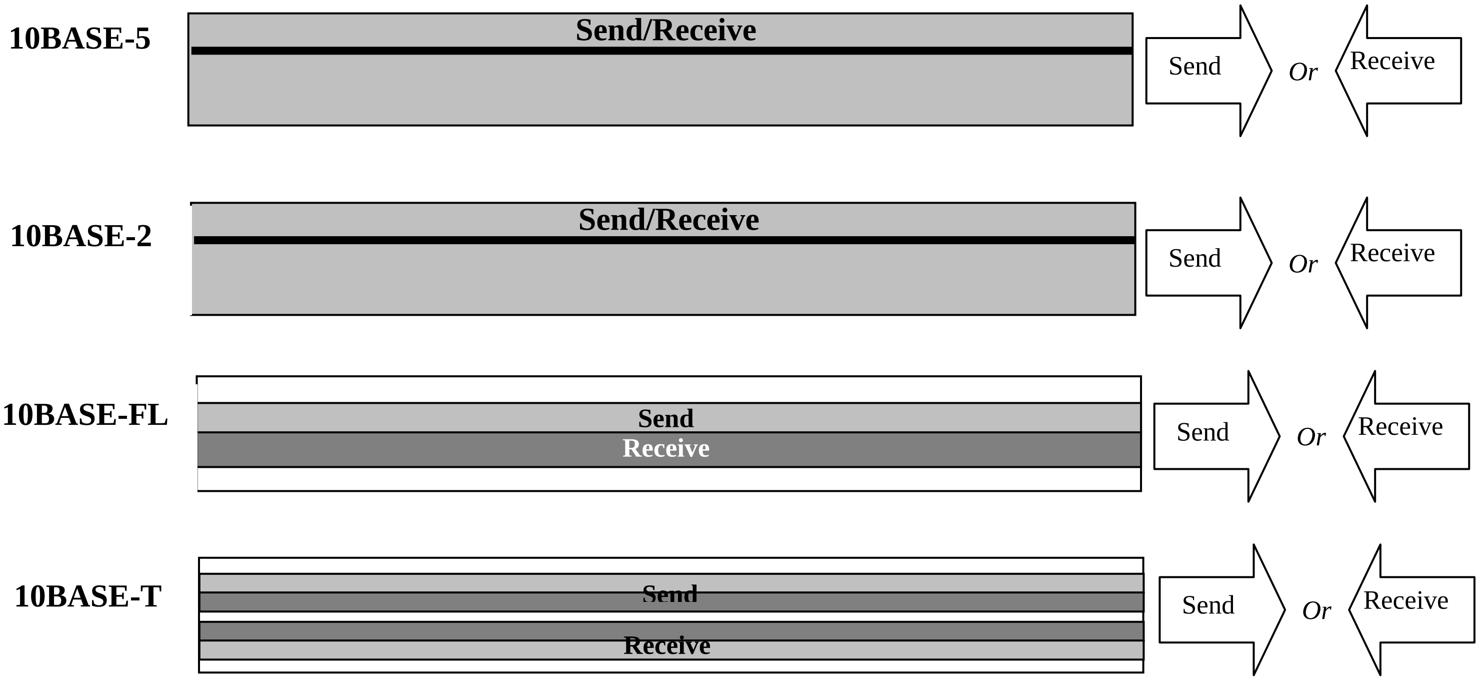


Figure 6: *Half/Full Duplex Support*

Forwarding techniques

Three different methods exist to switch a frame from one port to another. These are called forwarding techniques. Each technique has been optimised for use in a given situation.

Store and forward

This technique requires the whole of the frame (between 64 and 1500 bytes) to be received and “stored” in the switch before being forwarded. This gives the switch a chance to check that the frame is “good” (that it contains no collisions, it is a valid length and has a correct checksum) before propagating it through the network. The switch then uses its forwarding table to decide on which port to output the frame

Since each frame can be a variable length and the whole frame must be stored in the switch, delays are both variable and relatively long. This means that the store and forwarded method is best suited for use on a backbone network. On the downside, the variable delays mean that this method is not suitable for true multimedia applications, which require fast and consistent delay for sound and video to remain useable.

Store-and-forward must always be used when switching between connections of different speeds. For example, going from 10Mbps to 100Mbps.

Speed Conversion

This may be best described by an analogy of two people trying to jump from a moving slow train to a fast train whilst holding hands. If they don’t jump at the same time, the partnership breaks! The typical delay in this mode is between 54-1200 microseconds.

Delay in this mode will be around 50-2000 microseconds depending on the frame size.

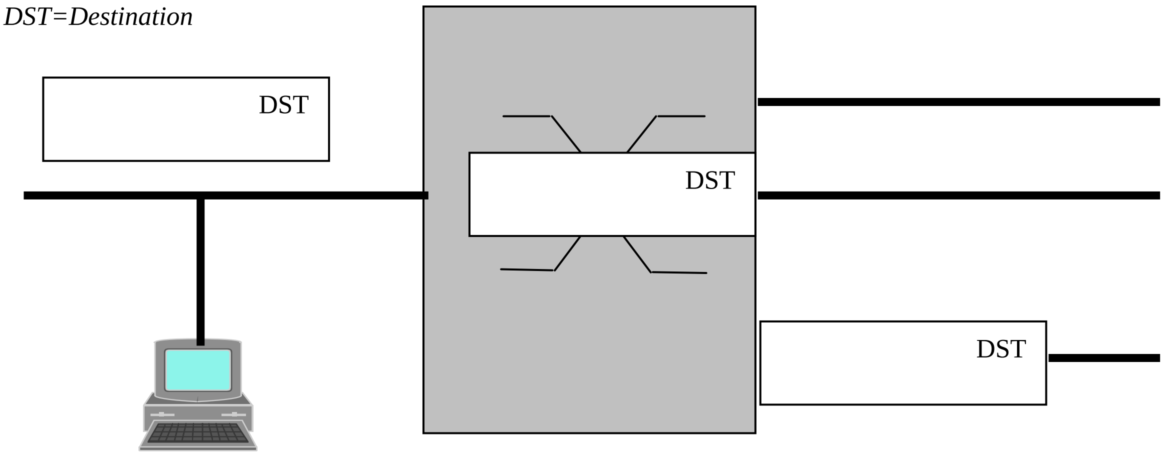


Figure 7: *In store and forward, the whole frame has to be stored in the switch before being forwarded*

Cut-through/Fast forward switching

This is the other extreme from store and forward. Here the switch waits until it has read the destination address and then begins sending the data through to the appropriate port. In this mode the data is already leaving the switch before it has finished coming in, offering a fixed delay.

This mode cannot be used when changing speeds in a switch. If the data entered at 10Mbps but left at 100Mbps, the input would be unable to keep up with the output.

A switch operating in cut-throat mode only looks at the destination address (**Figure 3**) of a frame. As soon as it has read the address, it starts sending the frame out of the appropriate port. The switch assumes the frame is valid and that is has had no collisions. If the frame has had a collision then the bad frame may be output to the receiving segments where it could cause collisions. The only way to guarantee that all forwarded frames are valid when using cut-throat is to have only one user per segment (then collisions will not occur). The typical delay in this mode is 20 microseconds.

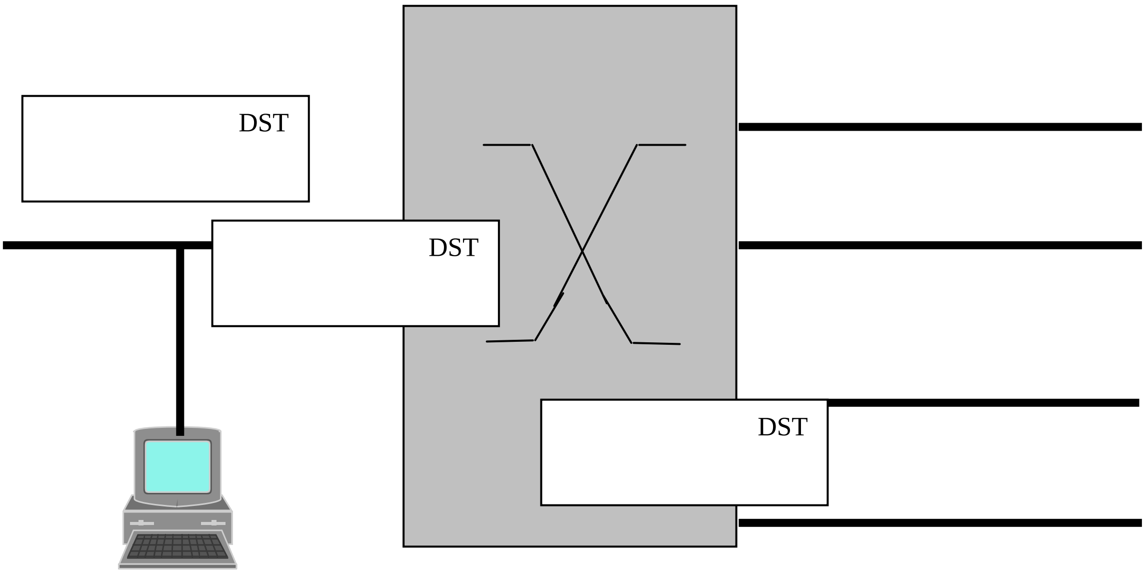


Figure 8: *In cut-through mode, the first 14 bytes (characters) need to be read by the switch before the frame is switched to the output port.*

Enhanced cut-through

There are two ways to modify the cut-through forwarding techniques.

For minimum latency, the switch could begin to broadcast all incoming frames on all output ports until the destination has been read. Then the switch continues sending on the correct destination port only. Unfortunately, this method creates a high amount of collisions everywhere on the network.

The second modification of cut-through forwarding allows for filtering. By reading each frame all the way to the protocol ID (LE field see **Figure 3**), the switch can then filter frames by source, destination address and protocol type, even in cut-through mode. This enables the Network Manager to maximise the security on the network while minimising latency in the switch.

Fragment free/Error free cut through

This is the most flexible of the three forwarding techniques. In this mode a switch will wait until it has received 64 bytes of data before starting to output the frame. Because of the way Ethernet functions (see the "How Ethernet Works" section), once 64 bytes have been sent, all other stations on the segment have had time to realised that the segment is busy and therefore hold back from sending.

Once the switch has seen 64 bytes it knows the frame is collision free and safe to forward onto the destination segment. Fragment free therefore offers the advantages of fixed delays (typically around 50 microseconds) and ensures only error free frames are forwarded, which ensures speed and security. The fixed delay makes fragment free the best forwarding technique to use on a backbone.

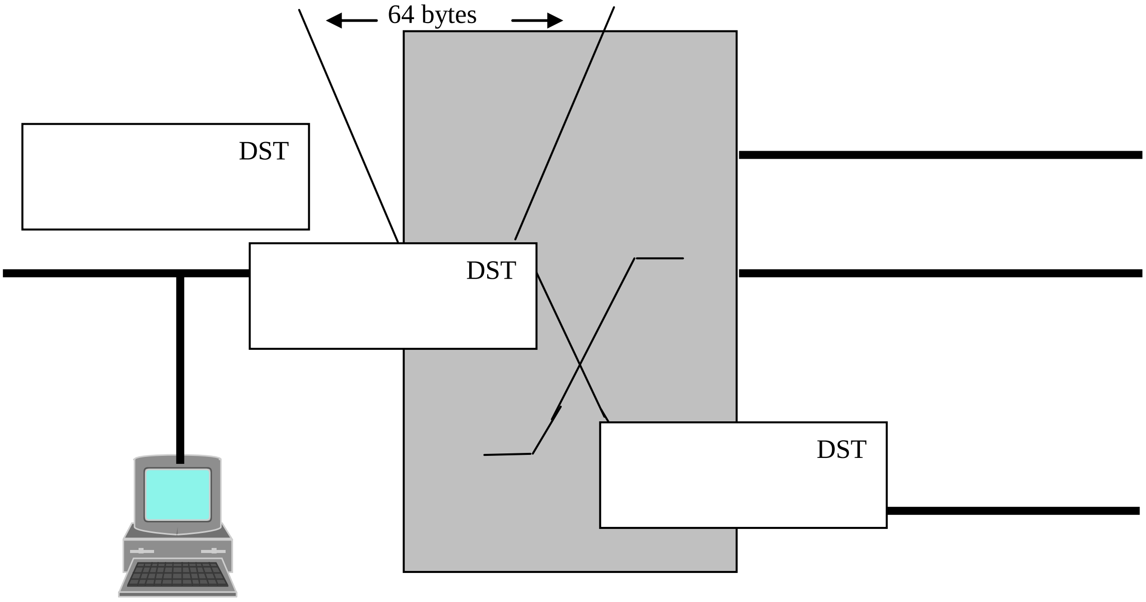


Figure 9: *In fragment free forwarding, the first 64 bytes are read by the switch before the frame is switched to the output port.*

Types of Switches

Several types of switch are now available with different costs and functionality. General terminology and when to use the different type of switches are explained below.

Backbone switches

Backbone switches (sometimes called enterprise switches) are implemented at the core of a network. They normally replace routers as they offer higher throughput rates, reduced delay and greater flexibility in terms of virtual LANs (described later). They must offer resilient PSUs and uplinks to Fast Ethernet, ATM or FDDI. They are normally modular and offer hot swap capabilities.

Enterprise switches must support segments with multiple users, so they must therefore support a high number of MAC addresses per port/switch. This means that they must contain a relatively large amount of RAM for storing address tables.

Since several Backbone switches can be connected together, they must support spanning tree to ensure that the network remains resilient in the case of failure.

A Backbone switch generally connects to a server, other switches, and in some instances to chassis or stackable hubs.

Since they connect to many different types of components, Backbone switches must be able to support different forwarding methods and duplex settings on a per-port basis. For example, a Backbone switch may need to connect with other switches and/or a file server using full-duplex and cut-through whilst also connecting to a hub using store-and-forward and half-duplex.

Workgroup switches

Workgroup switches support multiple MAC addresses on some or all ports (often configurable). They can have traditional hubs connected to each port, thus providing greater bandwidth for each user. Multiple 100Mbps ports may also be configurable. These can either be connected to file servers or provide an uplink to a Backbone switch.

To provide greater flexibility, these units may also have an ATM/FDDI link. Because several different types of connections may exist on a single switch, a choice of forwarding techniques must be available on a per-port basis.

Desktop switches

A Desktop switch is designed to replace a traditional hub. A Desktop switch supports a single workstation on each of the majority of its ports, therefore supplying each station with a private 10Mbps connection. This means that it need only support a single MAC address on most ports.

However, because one port connects to the "Backbone", the switch must have the ability to support multiple MAC addresses on at least one port. Just like other switches, Desktop switches should support a per port choice of forwarding techniques to allow for maximum flexibility. As these switches only support one station per port, cut-through forwarding is the optimum forwarding technique to use since collisions will not occur. However, fragment free or store and forward may sometimes be required for workgroup/backbone connections.

Usually, changing from a standard hub to a Desktop switch is fast and easy. The switch accepts the same cabling and connections as the hub it is replacing.

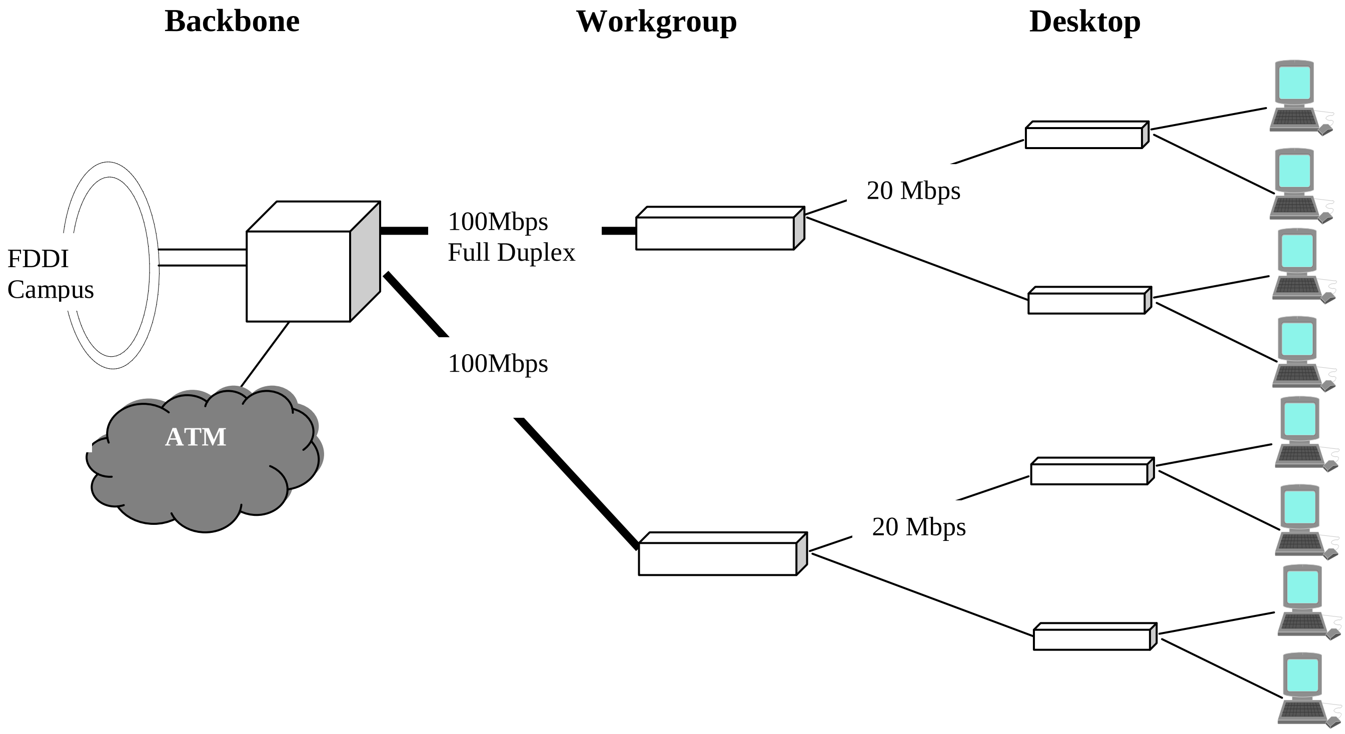


Figure 10: *Switches positioned*

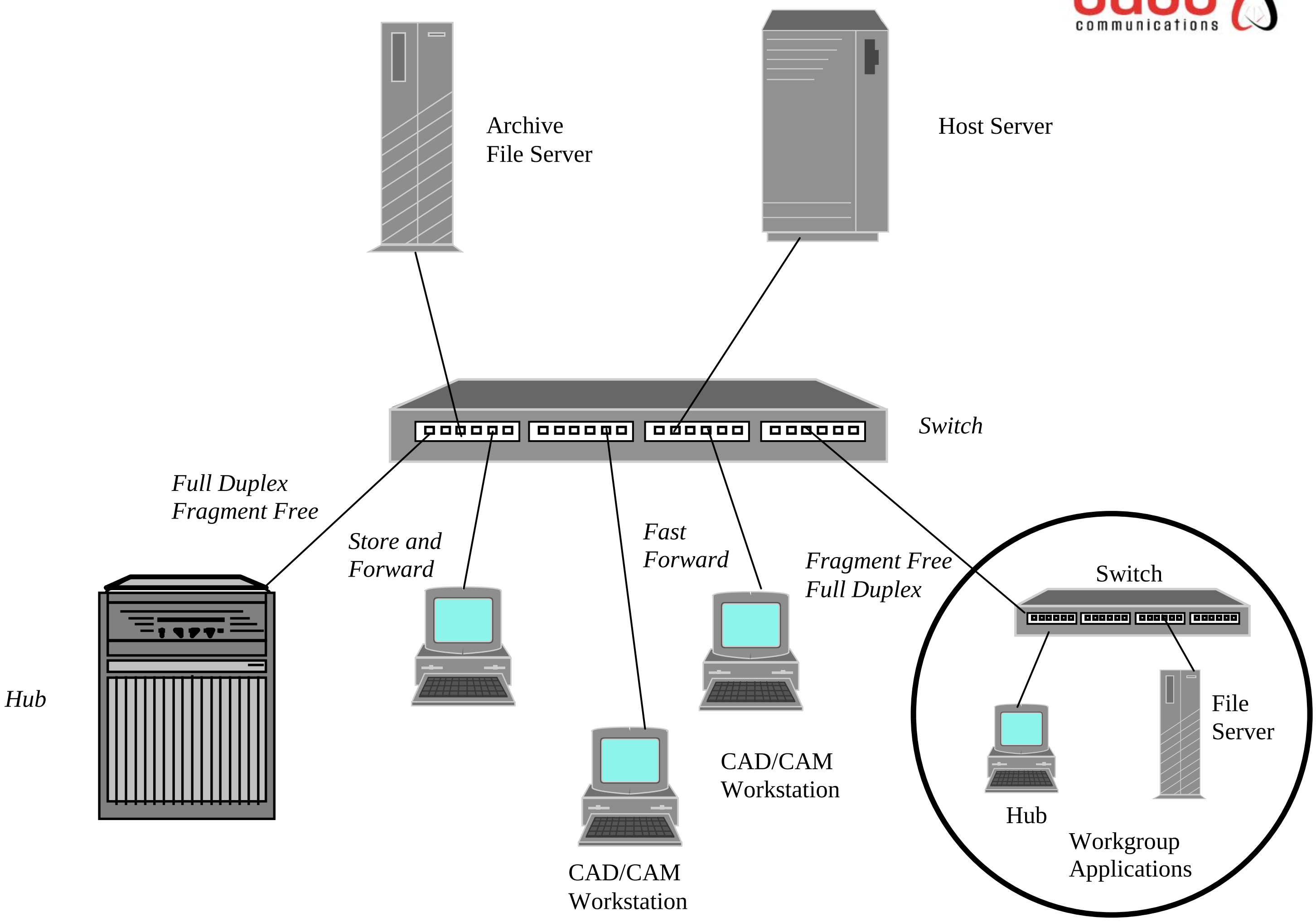


Figure 11: *Port by port configuration*

Traffic Control

Switching offers some unique methods of controlling traffic on a network. Flow control is required to enable the switch to continue to function when one or more ports become overloaded with traffic. The switch is also able to filter data by protocol type, source address, and/or destination address. Filtering increases security improves performance and makes the network simpler for users to navigate.

Flow control/Back pressure

Even with sophisticated buffering mechanisms it is still possible for more data to enter a switch than to leave it. This is where flow control is used. If the switch is running out of buffer it must find a way of stopping the source. One way of achieving this is by simulating a collision on the segment that is causing the problem (see **Figure 12**). This causes the transmitting device to stop and try to resend the transmission later – no data is lost. The switch will keep simulating collisions until the data in the switch's buffer is reduced.

Filtering

Filtering is a traffic control technique that prevents unwanted data from propagating on the network. It could also prevent users from accessing parts of the network for security reasons.

The switch can be configured, for example, to allow only users from the Accounting Department to access the server. Filtering can also be used to protect parts of the network from broadcast storms.

To filter the data, the switch must read the source, destination address and protocol of each frame. Therefore, filtering cannot be done in cut-through mode unless an enhanced cut-through technique is implemented.

Security

The need to have a secure network is paramount to protect sensitive information and to avoid both accidental and malicious damage. Since switches break up the network into independent segments, several sophisticated security options become available.

Each station is uniquely identified by its MAC address. A switch could be configured to allow communication between specified MAC addresses only. This ensures that users may only communicate with those servers and/or stations that they are specifically permitted to access – other stations will be invisible to unauthorised users.

Time-of-day access is a facility that can be configured on some switches to ensure that users only utilise the network at authorised times.

Protocol access security can also be added to ensure that only authorised protocols are used on each port.

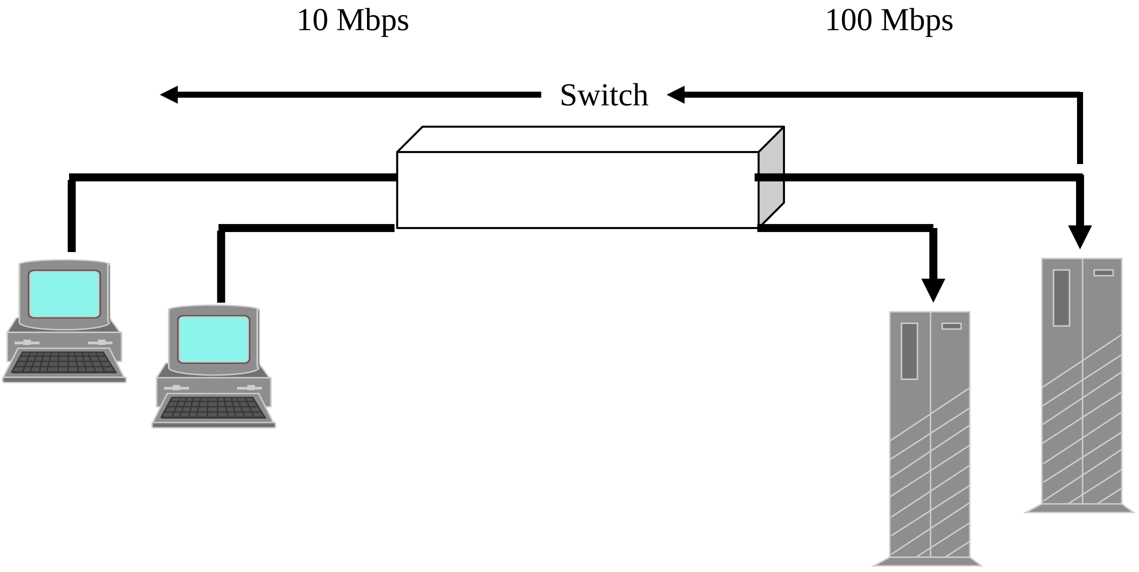


Figure 12: *Flow Control*

Price Performance

As explained in the "Brief History of LANs" section, most people judge a hub on its price per port and a router on the number of frames it can handle per second. Switches fit in between the two and therefore neither of the above is a good measure for judging switches.

Price performance takes both of these measures into account and is therefore the best way to measure a switch's performance. The original formula was devised by IBM's research division in Zurich and can be expressed as:

$$\text{Price/Performance} = \frac{\text{Cost per port x Number of ports}}{\text{Aggregate bandwidth x Efficiency}}$$

The value of the Efficiency parameter is determined by the switch architecture. It is the percentage of the network load that the switch can handle before no further throughout is possible.

Many switch manufacturers claim 100% efficiency, but this is only when sending a single frame through a switch with only two ports connected. Networks are never configured like this.

An example

An automobile manufacturer may claim a very rapid acceleration rate for a car. However, it assumes one driver and no passengers. If the car is going uphill and is loaded with people and baggage, then performance may suffer badly. It is the same with switches. The efficiency rating for a switch tells how the switch will perform under pressure.

As explained in the "Switch Hardware" section, output buffering combined with full ASIC design provides the most efficient switch architecture.

By incorporating efficiency into the equation, the price/performance ratio provides a true measure of a switch's performance – after all you don't want a switch to start dropping frames when you need it most – under load.

Management

The key problem is facing network managers today is how to manage their LAN installation proactively. Most managers first know of a problem when users start to complain about the response time on the network. By which time it is too late.

When a network is only connected via hubs, all traffic is seen at all p arts of the LAN. This makes management a simple matter of attaching a monitoring station somewhere in the network.

On the other hand, switches essentially create a new segment for each port. Traffic is sent only where it is needed for communication, so the place to monitor traffic is at the switch itself. It is very important that the switches you decide to use include built-in monitoring functions for each port.

SNMP

Simple Network Management Protocol (SNMP) is the de facto way to manage any switch. SNMP uses a series of closed questions to interrogate a device and provide answers. SNMP does not provide for graphical images of the product – this is a separate function provided by the management package.

SNMP utilises a significant amount of bandwidth to provide the raw information required by most management platforms.

With SNMP it is possible to have some idea about network traffic, but it is not easy to gain an accurate picture of what is going on and where potential problems could occur.

RMON

RMON utilises a management probe to collect information and then locally computes parameters such as utilisation and busiest host. RMON information is stored and passed to the management platform at regular intervals in the form of summarised reports. This substantially reduces the load on the network while providing a much higher level of information than SNMP.

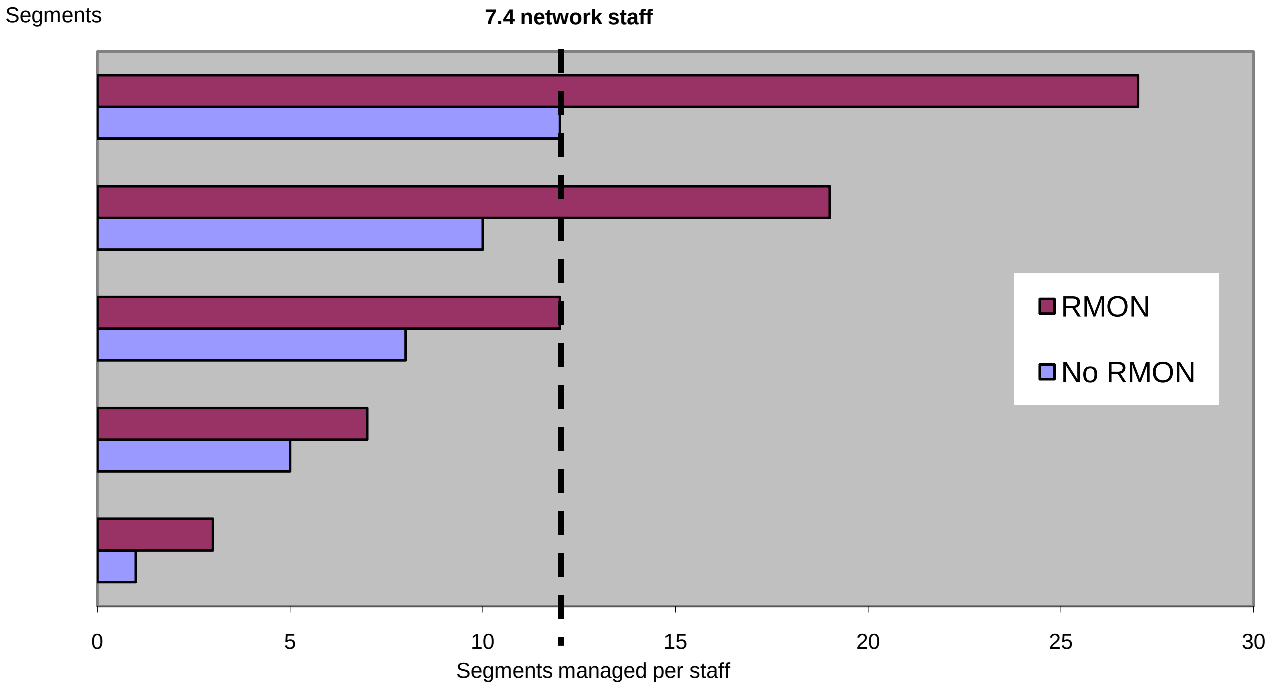
The RMON standard consists of nine groups of information (ten for Token Ring). The most important one is the Statistics Group, which samples the activity on the LAN at regular intervals determined by the manager. This is used by the History Group so that trend analysis can be made. This allows Network Managers to predict when problems will occur. The Host Group shows the busiest hosts on a segment. Other groups enable the RMON probe to generate alarms, capture frames, filter information etc.

These tools combine to enable a Network Manager to identify which stations need to be moved to a new segment in order to ensure the optimum throughput on the network. For example, over time, the History Group may show a gradual increase in utilisation on a crowded segment. This makes is possible to predict when it will hit the critical 40% utilisation level. The Alarm Group can identify this trend and notify the Network Manager. By examining the Host Group data, the Network Manager can then take action before a problem occurs.

The best way to monitor the many segments of a switched LAN is to make sure your switches include built-in RMON for each port. In this way it is possible to monitor the performance of all parts of your network proactively.

A switch designed with hardware-based RMON (usually implemented with ASICs) will be able to gather and process RMON statistics on each port without degrading switching performance. Statistics gathering and report generation will be done by the RMON ASIC independently of the switching functions.

Processor-based switches, on the other hand, will suffer a performance loss when RMON is applied. This is because monitoring will absorb CPU cycles that would otherwise be used to deliver switching functions.



Source: McConnell Consulting, 1994

Figure 13: *Diagram showing the impact RMON makes on managing a network*

Addressing

A fundamental function of a network switch is its ability to direct each frame to its correct destination. The switch must find out and remember where each station is located, and must make sure it does not create loops in the network.

Addresses per port

To direct traffic, a switch must know where to find each station in the network. It must learn and store (in an address table) which MAC addresses can be found at each port. Some switches allow a set number of MAC addresses to be stored per port. More sophisticated switches support a specific number of MAC addresses that can each be allocated dynamically per switch. The latter type is more flexible since is allows for more MAC addresses to be on some ports than others.

Passive mode

If working in a passive mode, a switch learns addresses as it sees them appear on its ports. Whenever it receives a frame destined for an address that it doesn't know, it generates a broadcast to all ports. The receiving station will reply and the switch will store the locations in its address table.

Active mode

In active mode, a switch requests the address of all devices on each segment. All stations respond so that the switch can update its address table. This is the fastest way for a switch to identify where all addresses are. This method keeps the switch informed about the network. It reduces the chance that an unknown address will appear, therefore reducing the number of broadcasts generated by the switch.

Address ageing

As time goes on, the switch will monitor the age of each entry in the forwarding table. If no activity is seen from an address after a specific amount of time (sometimes definable), the switch assumes the device has been switched off and removes the address from its table.

Spanning Tree Algorithm (STA)

With Ethernet it is important to prevent loops from appearing in the network. This would cause data to be regenerated and continue going round in circles, causing the Ethernet to be saturated.

The spanning tree algorithm (STA) prevents logical loops from appearing in the network. If multiple paths between two points exist, then the STA enables the switch to select the most efficient one. If a path goes down, STA finds the best way to work around the failure, thus sustaining network operation.

Spanning tree support is especially important in Backbone and Workgroup switches.

Virtual LANs

Virtual LANs will have a significant impact on networking. Today it is very difficult to support mobile users on a network. This is because a router's addressing scheme always expects to find a given "address" on a given segment. If that address moves to another segment, the router's filter tables need to manually changed. This is a very time-consuming job, so user mobility is not normally encouraged in a router-based network.

With switches the story is very different – switches use MAC addresses to make decisions on "who connects to who" – they can therefore build up virtual teams of users based on MAC addresses. Regardless of location, the switch will still allow users of predefined workgroups to communicate.

Virtual LANs are also important for implementing security and preventing broadcast storms.

Virtual LANs can be supported in two ways: by physical switch port or MAC Address.

Port based

Port based virtual LANs are organised by physical port number. For example, switch ports 1, 4, 7 and 9 could be one virtual LAN with ports 3, 6, 10 and 12 being another. LAN broadcasts from servers within each group would only go to other members of its virtual LAN. This ensures that broadcast storms cannot cause a network meltdown through volumes of traffic.

MAC address based

As the name implies, this type of virtual LAN is based on the MAC address of the station. This means that a user with a portable PC can connect it at any point on the network. The switch will ensure that the user is connected to the correct workgroup.

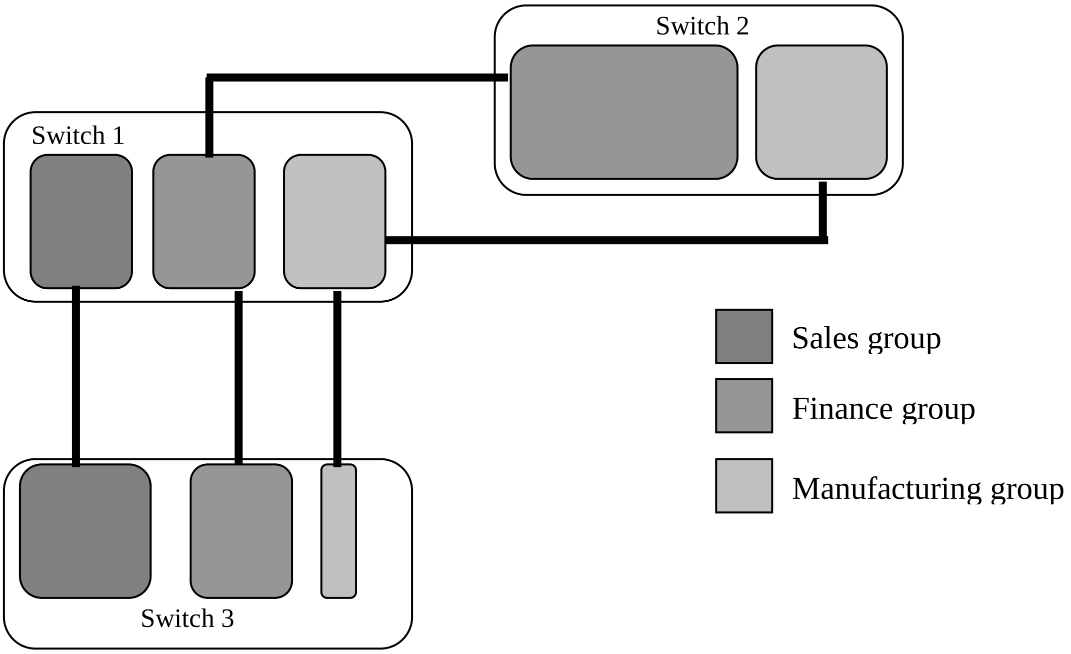


Figure 14: *Virtual LAN Support*

New backbone structures

Fast Ethernet

AS the name implies, Fast Ethernet works faster than the 10Mbps Ethernet. In fact it works at a speed of 100Mbps. Fast Ethernet can run over other existing 10Mbps cable.

100Mbps Ethernet can also run over fibre, which is ideal for connecting Backbone switches to Desktop switches over longer distances. The best place to utilise fast Ethernet is between switches running full-duplex (giving 200Mbps capacity) and when connecting servers to switches.

Desktop switching enables clients to work at 10Mbps (20Mbps if there is one client per port working full-duplex) while the file server connection runs at 200Mbps. This eliminates potential bottlenecks by providing extra bandwidth for high-traffic switch/server and switch/switch connections.

ATM

ATM uses fixed-length frames, whereas Ethernet allows variable length frames (from 64 bytes to 1500 bytes). Ethernet therefore requires a high level of processing at the receiving end, so a station will know when it has received the entire frame.

ATM uses a technique called cell switching, which uses fixed-length frames. Each cell is 48 bytes (or characters) long with 5 bytes of overhead. Since each cell is uniform, processing can be done quickly and efficiently by hardware. With volume, this will make ATM solutions cheaper. However, only after the standard are all agreed upon (by both the LAN and WAN vendors) and silicon-based ATM devices start to become available will ATM become practical. Because of the high overhead, ATM is only suitable for high speed transmission (25Mbps and above).

Other Types of Switching

Token Ring switching

Token Ring, unlike Ethernet, is deterministic. Therefore the bandwidth is much more predictable. However, switching on a per port basis rather than sharing the Ring as usual will increase the bandwidth in a similar way to Ethernet switching.

FDDI switching

FDDI (Fibre Distributed Data Interface) uses a circular cable structure, which is like Token Ring, only faster. Originally, FDDI was a fibre optic standard, but today both fibre and copper wire versions exist. Since it is a fast standard (100Mbps) but also relatively expensive, it most often used for LAN backbones.

In the short and medium term, FDDI is one of the best backbone technologies to use because it is proven, resilient, and maintain its performance under a heavy load.

Port switching

This is not Ethernet switching as such but it can act on an Ethernet network. Port switching normally switches between backplane segments within a hub. For example, Port 1 could connect to Backplane 3 and Port 2 to Backplane 1. The net result is that several users still end up sharing the same backplane creating, in effect, a collapsed backplane, but with greater levels of flexibility.

Where to use switching

In the main switching exists as a way of increasing bandwidth and reducing delay on existing networks. If your network is becoming congested, then replacing your hubs with switches will go a long way towards solving your problem.

An Ethernet switch will easily replace an existing hub, using all of the existing cables, PC network cards, and backbone connections. Only the hub must be replaces, and in many cases could be reused in combination with switches.

New Networks

If you are designing a new network, use switches instead of hubs to connect to each Desktop and to the Server or Backbone. This will prevent traffic problems and enable your network to support true multimedia applications in the future.

Switching also enables each workgroup to be connected to workgroups and servers on other floors via a 100Mbps backbone. Your investment in switches will be protected because they prepare your network for easy upgrade to an ATM backbone.

Heavily loaded networks

According to network theory (a full explanation of which is beyond the scope of this book), a typical shared Ethernet network reaches maximum throughput when it is utilised less than 37% of the time. After that, throughput actually goes down due to collisions.

If utilisation on your network is higher than this limit (due to a large number of stations, or a few very active users), then you are probably already having problems with congestion. Converting some or all of your hubs to Ethernet switches will dramatically improve performance.

Since switches allow single station segments, collisions can be eliminated altogether, enabling utilisation up to 100%.

If only a few stations require extra bandwidth (such as servers or graphic stations), then you should connect each of these to dedicated ports or Fast Ethernet ports. A less demanding group of users can then be connected via a shared hub to a single port on the switch. In this way, you will conserve switch ports and still meet the needs of every user.

Uplinking to fast Ethernet or ATM

As a rule, 10Mbps is the most a user's workstation will need for the foreseeable future. However, connections between two switches or from a switch to a file server may need to support several conversations, each being switched to a station at 10Mbps. To avoid bottlenecks, these connections should be at a higher bandwidth.

Most switches solve this problem by including one or more ports that can connect to 100Mbps Fast Ethernet. Others include expansion slots for FDDI or ATM connections.

Connections between switches can often be full-duplex, which doubles the effective bandwidth to 200Mbps for Fast Ethernet or 20Mbps for standard Ethernet.

Multimedia applications

Multimedia applications, such as interactive training and video conferencing, require a large amount of bandwidth. They are also very sensitive to variable delays, which are an unavoidable consequence of shared Ethernet.

By connecting multimedia users directly to a switched port instead of a shared hub port. Bandwidth will jump to 10Mbps and collisions will never occur. When used in conjunction with fragment free forwarding, true multimedia will be possible on Ethernet.

Glossary

10BASE-2

A form of Ethernet and IEEE 802.3 network cabling, using thin co-axial cable. The term refers to 10Mbps (speed) BASEband (transmission) 200 metres (maximum length actually 185 metres). Commonly known as “Cheapernet”.

10BASE-5

A form of Ethernet and IEEE 802.3 network cabling, using thick co-axial cable. The refers to 10Mbps (speed) BASEband (transmission) 500 metres (maximum length).

10BASE-T

A form of Ethernet and IEEE 802.3 network cabling, using twisted-pair cabling. The term refers to 10Mbps (speed) BASEband (transmission) twisted-pair cable. 10 BASE T recommends a segment length of 100 metres.

ANSI

American National Standards Institute.

AppleTalk

A proprietary network system designed by Apple Computer, principally used with Macintosh systems.

Application Layer

OSI standard for reading, writing, sharing and administering files on different machines.

ARCnet

A proprietary network system designed by Data-point Corporation using token Bus techniques.

ARPANET

The Defence Advanced Research Projects Agency Network. A network in America which was used to develop the Internet Protocol Suite, source of TCP/IP protocols and many others such as OSI.

ASCII

American Standard Code for Information Interchange

ATM

Asynchronous Transfer Mode, a high-speed, scaleable cell-switching protocol that breaks packets down into fixed 53-byte cells.

AUI

Attachment Unit Interface. Defined in IEEE 802.3 as the interface between the transceiver and the network device.

Back Pressure

A technique for preventing buffer overload by sending a jam signal when a station tries to send to an overloaded port. The sending station will interpret the signal as a collisions, stop sending and try again later.

Baseband

A transmission technique employed in LANs. Only one device can communicate at one time and a device simply sends its digital signal on to the cable without any modification.

Bridge

A device which connect two separate networks at the MAC layer and passes data between the two networks, filtering local communications.

Broadband

Analogue transmission technique employed in LANs. The signals on the network are divided, (usually by frequency division), to allow more than one (usually pairs of signals) on the cable at any one time. This form of communication requires modems for all devices wishing to access the network.

Bus

A form of network structure based on a single length of cable to which all network devices are connected. Ethernet is an example of such a network structure.

Bit

An abbreviated form of Binary digit used to represent a single transition on a data line, or a fraction of a byte.

Byte

A unit of computer information, consisting of eight bits (binary digits).

CAD

Computer Aided Design.

CCITT

Comite Consultatif Internationale de Telegraphique et Telephonique. A standards body concentrating on the definition of European Communications Standards.

Cheapernet

A term generated following a revised specification for Ethernet using a thin, lower cost, co-axial cable.

CLNS

ConnectionLess-mode. Networking Service as defined in ISO 8473, operating at level 3, Network, of the OSI reference model.

Co-axial cable

A cable consisting of a single inner core surrounded by insulators and a stranded metal sheath, all encased in a single insulating outer cover.

Comite Consultatif Internationale de Telegraphique et Telephonique

See CCITT

CONS

Connection-mode Network Service. Operates at the Network Layer (layer 3) and implements a connection-oriented networking system using X.25.

Connection Oriented Protocol

A form of transmission where a connection is established across a network proper to the first transmission of data taking place. The connection handles checking and error recovery during transmission thus incurring a connection overhead on the transmission.

ConnectionLess Network Service

See CLNS

Connectionless Protocol

A form of transmission where no communication takes place prior to the first transmission of data. Therefore no connection is made on the network and no checking takes place during transmission thus reducing the connection overhead on the transmission.

CSMA/CA

Carrier Sense Multiple Access with Collision Avoidance. A method of network access not covered by OSI standards. This technique is implemented on AppleTalk networks.

CSMA/CD

Carrier Sense Multiple Access with Collision Detect. The technical reference to the standard LAN transmission network utilising a bus structure, conforming to IEEE 802.3 and ISO 8802-3 which were based on the Ethernet network standard.

Cut-through Switching

A switch forwarding method designed to minimise delay. A packet is forwarded to the output port as soon as the destination address is known.

Data Link

Layer 2 of the OSI 7-layer model.

Datagram

A unit or packet of data sent across a network. A datagram is self contained, with source and destination address, but is not necessarily a complete communication, which may require many such datagrams.

DDCMP

Digital Data Communications Message Protocol. A form of transmission for wide area connectivity over DECnet.

DECnet

A network architecture produced by Digital Equipment Corporation encompassing local and wide area connectivity.

Distributed Systems Gateway

DSG. A device which operates at layer 4 of the OSI reference model and allows interconnection of dissimilar networks.

DIX DEC, Intel and Xerox

These terms are used in relation to the published Ethernet Specifications.

Early Token Release

A modified version of token passing where a device does not have to wait for its communication to circle a network before transmitting the token on around the network.

Ethernet

A LAN transmission network utilising a bus structure, first produced by Xerox in the USA, later adopted by both DEC and Intel as well (see DIX). This was later adapted to create the IEEE 802.3 and ISO 8802-3 standards (see CSMA/CD).

EtherTalk

An Ethernet implementation of the lower two layers of the AppleTalk protocol suite. EtherTalk complies with 802.3.

FDDI

Fibre Distributed Data Interface. A token passing ring network specification developed by ANSI, implementing dual optical fibre rings. The network operates at 100Mbps and can be up to 100 kilometres in length.

Fibre Distributed Data Interface

See FDDI

Fibre Optic Cable

A cable type based on a continuous strand of very fine glass which uses light to carry information. This type of cable is able to carry very high levels of data for great distances without interference.

File Server

A device, usually a PC, offering on-line storage and management of files for more than one device on a network. File servers offer storage and recall of data files and applications. Additionally they may offer sharing of files and applications, protocol conversion and central facilities such as electronic mail.

File Transfer Protocol

See FTP.

Fragment Free

A switch forwarding method that offers fixed delay whilst ensuring only valid frames are forwarded.

FTAM

File Transfer, Access and Management

FTP

File Transfer Protocol. A Protocol from the Internet Protocol Suite which provides transfer of files between two dissimilar machines.

Full-duplex

Two-way communication in which a station can send and receive data simultaneously, therefore doubling the bandwidth.

Gateway

A device used to allow conversion on LAN networks. Normally operating a layer 4 or above in OSI terms. Gateways convert information between two different networks, computers or applications.

Half-duplex

A communication method in which a station can either send or receive data, not both at the same time.

Half Repeater

A device which extends the distance a LAN can cover by joining two lengths of cable over another communication medium, such as a dial-up circuit, and regenerating the signal.

Head End

A device used with Broadband networks. This is positioned at one end of the cable to convert the signal from one channel and output it on the other channel.

Host Servers

A device which connects to a LAN and then allows a computer, which cannot directly support LAN protocols, to connect to it providing all necessary LAN support.

IEEE

Institute of Electrical and Electronic Engineers. An American Institute responsible for developing and publishing many communications standards (see section listing Standards).

IMAGE

A file format used within the Internet Protocol Suite for handling graphics

Institute of Electrical and Electronic Engineers

See IEEE

Integrated Services Digital Network

See ISDN

Interworking

A term extending beyond communication any relating to hardware and software compatibility not just to allow connection but to maximise operational efficiency.

International Standards Organisation

A suite of network protocols originally designed for ARPANET in America and since adopted as the main de facto protocols for LANs.

ISDN

A form of public network that is designed to handle voice and data at high speed in digital format.

ISO

International Standards Organisation. An organisation devised to create standards which will promote internetworking between different vendors equipment. The main focus of communications is implementation of the Open Systems Interconnection (OSI) reference model.

LAN

See Local Area Network.

LAN Manager

A networking communication system operating on OS/2-based PCs, plus other systems produced by Microsoft and 3Com.

Laser Printer

A high quality printing device capable of producing totally free format image output.

Latency

The time delay introduced by a network device. The time taken for the device to receive a frame and begin sending it out again.

LF

Line field used to identify the protocol in use.

LLC

Logical Link Control. The upper of the two sub-layers in layer 2 of the OSI 7 layer model.

Local Area Network

A system for intercommunication between computer terminals, PCs and related equipment operating within the same general area.

LocalTalk

A twisted pair transmission system used with Macintosh networks, designed by Apple Computer.

Logical Link Control

See LLC.

Logical Ring

When token passing is implemented on a Token Bus network, the token cannot circulate as on a Token Ring network. A Logical Ring is implemented where each device is given a sequential address. When a taken is relayed on the network each receiving address attaches the next logical device address thus ensuring circulation of the data.

MAC

Media Access Control. The lower of the two sub-layers in layer 2 of the OSI 7 layer model.

MAN

Metropolitan Area Network. A term often used to describe the latest, high speed and long distance LANs, such as FDDI.

MAP

Manufacturing Automation Protocol. An OCI initiative pioneered by General Motors to create a series of standards for computer and communication s equipment interworking on the production floor, for example central computer to production line robot.

MAU

Multi-station Access Unit. A device for use on IBM Token Ring networks that allows terminals, PCs, printer and other devices to be connected in a star-based configuration.

Media Access Control

See MAC.

Metropolitan Area Network

See MAN.

MMFS

Manufacturing Message Format Standard. An OSI Application Layer standard devised for MAP type networks.

Multi-station Access Unit

See MAU.

Netbios

A network communication system operating on MS/DOS PC networks.

Netware

A suite of application oriented interface protocols designed for use on LANs and produced by Novell.

Network Layer

Layer 3 of the OSI reference model which allows interconnection of dissimilar networks (see Distributed Systems Gateway).

Non-deterministic

A term relating to the inability of a designed being able to predict the performance or delay on a network where collision is possible.

Open Systems Interconnection

See OSI

OSI

Open Systems Interconnection. The term defined by the International Standards Organisation as a basis for standards to enable different vendor systems to interwork without modification.

PC

Personal Computer.

PDU

Protocol Data Unit. Protocol Data Units are the datagrams relating to Logical Link Control (LLC) transmissions.

Physical Layer

Layer 1 of the OSI reference model Presentation Layer, Layer 6 of the OSI reference model Probabilistic. A term relating to the ability of any one device being able to communicate on a network where collision is possible.

Protocol Data Unit

See PDU.

PSU

Power Supply Unit.

RAM

Random access memory used to temporarily store data in a CPU.

Rat's Tail

A term used to describe the network adaptor used to connect Apple Macintoshes to a LocalTalk LAN.

Repeater

A device which extends the distance a LAN can cover by joining two lengths of cable and regenerating the signal.

RG58

A co-axial cable, often referred to as “Thin” co-ax. This was introduced to reduce the costs of cabling early networks, particularly Ethernet. This version became known as “Cheapernet”.

Router

A device which connects two separate networks at the Network Layer. It operates in a similar way to a bridge but also has the ability to choose routers through a network. Because a router operates at the Network Layer it is protocol-dependent.

Segment

In bus structure network this refers to a single length of cable and attached devices which would normally be linked to another segment by a repeater.

Session Layer

Layer 5 of the OSI reference model.

Simple Mail

See SMTP

Store-and-Forward

A forwarding technique designed to avoid propagating bad frames. The frame is full loaded and checked before any of it is sent out again.

Transfer Protocol SMTP

Simple Mail Transfer Protocol.
A protocol from the Internet Protocol Suite which provides electronic mail transfer across a network.

SNA

Systems Network Architecture. An architecture and suite of protocols designed and produced by IBM for hierarchical computer networks.

Structured Wiring

A cabling philosophy which endeavours to minimise cable changes once a building has been wired.

TAP

A connection into the Local Area Network cable.

TCP/IP

Transmission Control Protocol/ Internet Protocol. Two main network communication protocols, part of the Internet Protocol Suite.

Technical and Office Protocols

See TOP.

TELNET

A protocol from the Internet Protocol Suite which provides a Virtual Terminal interface for communicating devices.

Terminal Servers

A device which connects to a LAN and then allows one or more terminals, which cannot directly support LAN protocols, to connect to it providing all necessary LAN support.

Token

A token is a special sequence of data which is passed around a network sequentially. When the token is captured by a device it flags it as busy and adds data to the frame. Once the transmission has circled the network the busy token is replaced by a free token which the next device can then capture and transmit.

Token Bus

A LAN transmission network utilising a bus structure and implementing token passing access. This network is detailed by the IEEE 802.4 and ISO 8802-4 standards.

Token Ring

A LAN transmission network utilising a ring structure. The most common form is IBM Token Ring. This network type is detailed by the IEEE 802.5 and ISO 8802-5 standards.

TokenTalk

A Token Ring implementation of the lower two layers of the AppleTalk protocol suite.

TOP

Technical and Office Protocols. An OSI initiative pioneered by Boeing Computer Services to create a series of standards for computer and communications equipment interworking in the office environment.

Transmission Control Protocol/Internet Protocol

See TCP/IP.

Transport Layer

Layer 4 of the OSI reference model.

Twisted Pair Cable

The most common form of communication cabling used by telephones, computer terminals and LANs. A pair of wires, coated in a plastic sheath are twisted together. In most cases several pairs are then bound in a single sheath and then laid as a single cable using multiple twisted pairs.

VDU

Visual Display Unit.

Video Conferencing

A form of communication requiring very high capacity networks, which transmits pictures of all parties as well as voice.

Virtual Network

A form of network where communication and access is a achieved without any knowledge of the network structure or location of a specific resource.

WAN

Wide Area Network. A network which uses public or private circuits to link terminals, PCs and computers over long distances that also often allows some user choice in destination. Followings the implementation of routers, bridges and gateways many LANs are becoming Wide Area Networks, but more typically networks such as X.25 are classed as WANs.

X.400

An Application Layer OSI standard for transferring electronic mail on a store and forward basis between different machines.

XNS

Xerox Networking Systems. An Ethernet network produced by the Xerox Corporation.

Yellow cable

A co-axial cable, often referred to as “Thick” co-ax. This was the first cable used on may early LANs.