CHAPTER 4

SWITCHING FUNDAMENTALS AND NETWORK APPLICATION LAYER PROTOCOL

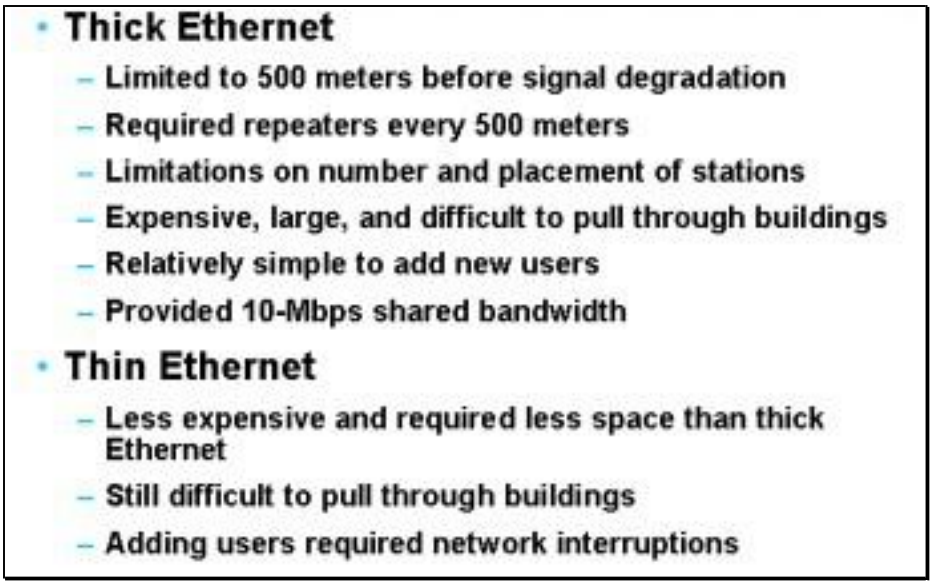
Switching Fundamentals

Switching is a technology that decreases congestion in local-area networks (LANs) by reducing traffic and increasing bandwidth.

Shared LAN Technology

4.1.1 Early Local-AreaNetworks

Figure 1: Early Local-Area Networks



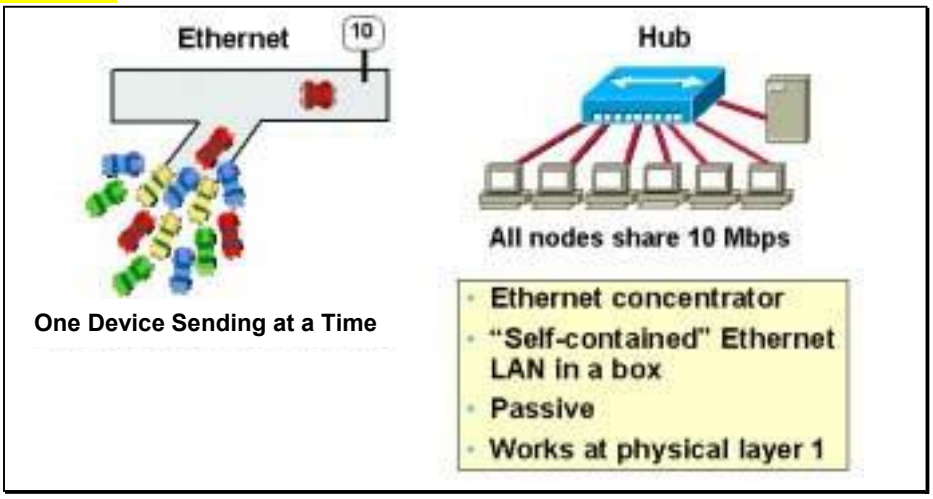
The earliest local-area network (LAN) technologies that were installed widely were either thick Ethernet or thin Ethernet infrastructures. It's important to understand some of the limitations of these infrastructures to see where LAN switching stands today. Thick Ethernet installations had some important limitations, such as distance, for example. Early thick Ethernet networks were limited to only 500 meters before the signal degraded. For distances beyond 500 meters, repeaters were required to boost and amplify that signal.

There were also limitations on the number of stations and servers on a network, as well as the placement of those workstations on the network. The cable itself was relatively expensive, and it was large in diameter, making it difficult to install throughout the building because it was pulled through the walls and ceilings. Adding new users was relatively simple—a nonintrusive tap plugged in a new station anywhere along the cable. The thick Ethernet network provided a capacity of 10 megabits per second (Mbps), but this bandwidth was shared, meaning that 10 Mb was shared among all users on a given segment.

A slight improvement to thick Ethernet was thin Ethernet technology, commonly referred to as *cheaper net*. This technology was less expensive, and it required less space in terms of installation than thick Ethernet because it was thinner in diameter, hence the name. It was still relatively challenging to install, though, because it sometimes required a *home run*, or a direct run from a workstation back to a hub or concentrator. Adding users required a momentary interruption in the network, because a cable segment had to be broken in order to add a new server or workstation.

4.1.2 Hubs

Figure 1: Hub Addressed Many of These Problems

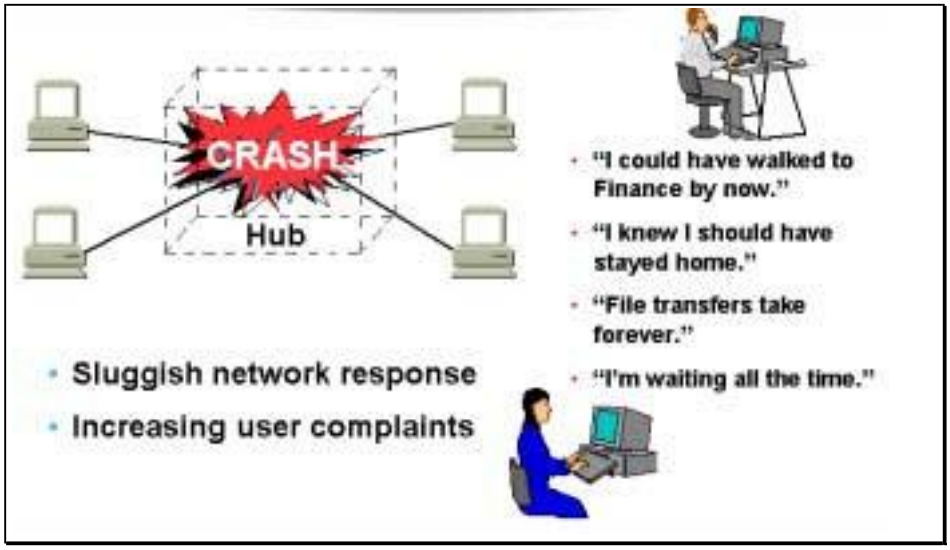


Adding hubs or concentrators into the network offered an improvement on thin and thick Ethernet technology. Hubs are sometimes referred to as Ethernet concentrators or Ethernet repeaters; they are basically self-contained Ethernet segments within a box. Unshielded twisted-pair (UTP) cabling was used, but the fundamental limitation of a shared technology remained. As you can see in Figure [1], Ethernet is fundamentally a *shared* technology—all users of a given LAN segment "fight" for the same amount of bandwidth. This situation is analogous to cars all trying to get onto the freeway at once. In the network, even though each device has its own cable segment that connects into the hub, they all share the same fixed amount of bandwidth. Frames, or packets, in a network all vie for bandwidth.

Although physically it looks like all users have their own segment to their workstation, they are all interconnected inside the hub, so the hub is still a shared Ethernet technology. Also, these devices are passive, meaning that they're virtually transparent to the end users—the end users don't even know that they exist. In addition, the devices have no role in terms of a forwarding decision in the network, nor do they provide any segmentation within the network because they work at Layer 1 in the OSI framework.

4.1.3 Collisions

Figure 1: Collisions: Telltale Signs

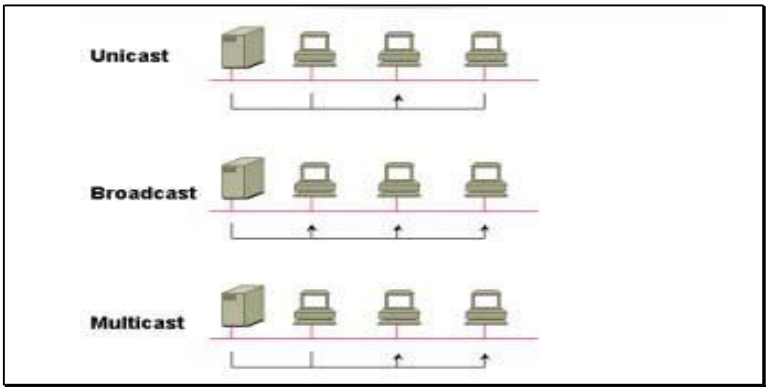


Collisions are by-products of an Ethernet network. In an Ethernet network, many stations share the same segment, so any one of these stations can transmit at any given time. If two or more stations try to transmit at the same time, a collision results, indicating that the network is becoming too congested or that too many users are on the same segment.

When the number of collisions in the network becomes excessive, sluggish network response times result; an increasing number of user complaints reported to the network manager is a good indication that the network is sluggish.

4.1.4 Transmission Ways

Figure 1: Other Bandwidth Consumers



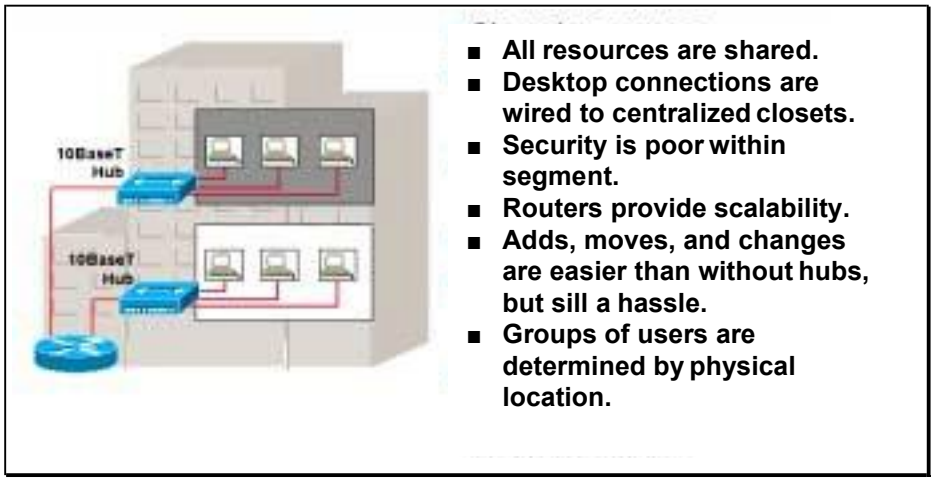
It's also important to understand fundamentally how transmissions can occur in the network. Communication in a network occurs in three ways. The most common way is by *unicast transmissions*. In a unicast transmission, one transmitter tries to reach one receiver. This form of communication is by far the most common form of communication in a network.

Another way to communicate is by *broadcasting*, when one transmitter tries to reach all receivers in the network. As you can see in Figure **[1]**, the server station is sending out one message and it is being received by everyone on that segment.

The last mechanism is known as a *multicast*, when one transmitter tries to reach not everyone, but a subset or a group of the entire segment. As shown in Figure **[1]**, two stations are reached, but one of them doesn't need to participate, so it is not in the multicast group.

4.1.5 Hub-Based LANs

Figure 1: Hub-Based LANs

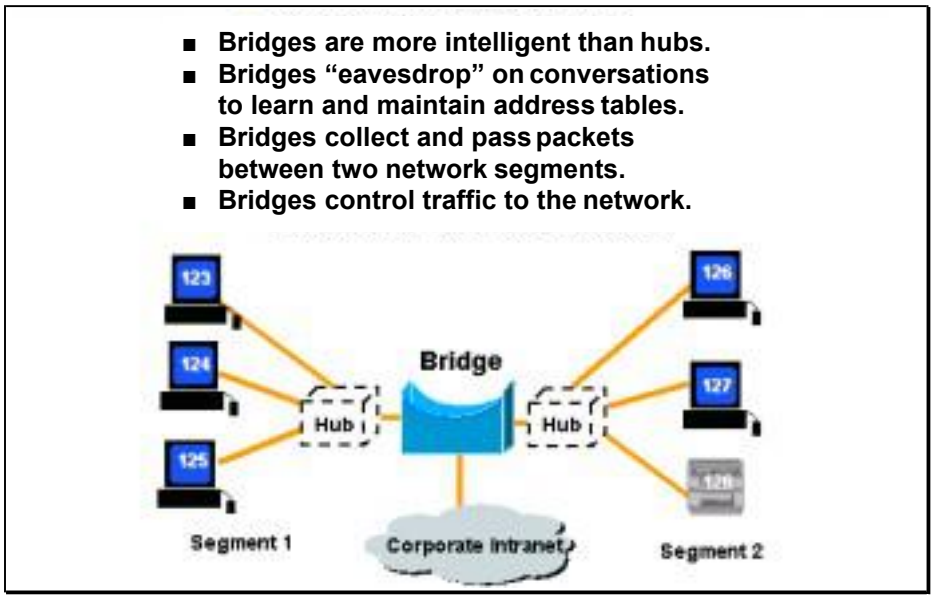


Hubs were introduced into the network as a better way to scale thin and thick Ethernet networks. It's important to remember, though, that these are still shared Ethernet networks, even though hubs are used. Each individual workstation or server in the network has an individual desktop connection, allowing centralization of all cabling back to a wiring closet. This setup makes adds, moves, and changes easier because cables can just be moved around in the wiring closet.

In a hub- or concentrator- based network, workgroups are determined simply by the physical hub plugged into. (Again LAN switching makes configuration for workgroups even easier.)

4.1.6 Bridges

Figure 1: Bridges

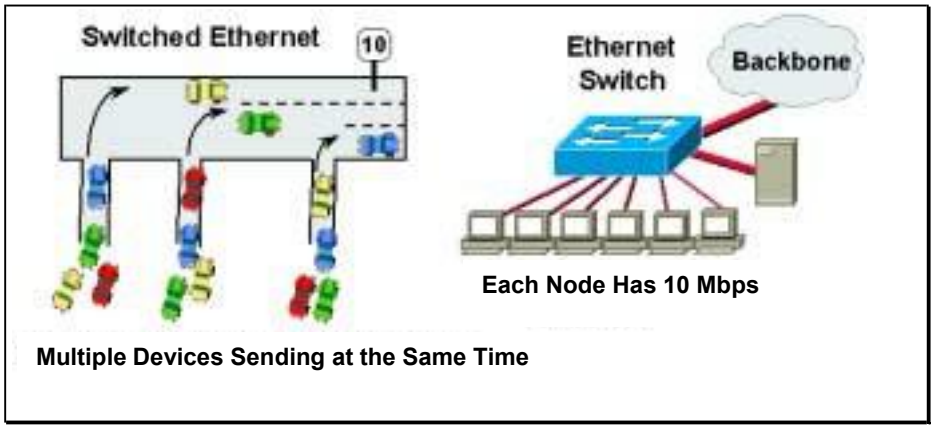


Segmentation is used to scale networks. One way to scale hub-based networks is to add routers; another is to add bridges, which provide a certain level of segmentation by adding a certain amount of intelligence into the network.

Bridges operate at Layer 2, whereas hubs operate at Layer 1. Operating at Layer 2 offers more intelligence for making forwarding decisions. Bridges are more intelligent than hubs because they can actually listen in, or "eavesdrop" on the traffic going through—they can look at source and destination addresses, and they can build a table that enables them to make intelligent forwarding decisions. They actually collect and pass frames between two network segments while at the same time making intelligent forwarding decisions. As a result, brides can provide greater control of the traffic within a network.

4.1.7 Switches—Layer 2

Figure 1: Switches—Layer 2



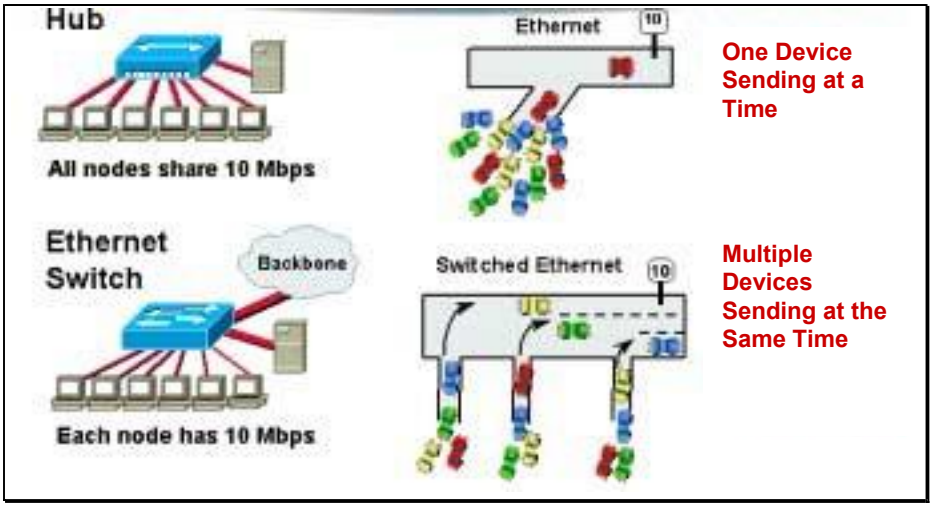
Switches provide even better control, at least at Layer 2. As you can see in Figure [1], we've improved the model of traffic going through the network. Returning to the traffic analogy, we have actually subdivided the main highway so that each car has its own lane to drive on through the network. And fundamentally, this functionality can be provided in data networks as well.

Each station has its own cable into the network—analogous to each workstation having its own "lane" through the "highway." Then with functionality known as *microsegmentation*, each workstation gets its own dedicated segment through the network. Microsegmentation enables the travel of multiple, simultaneous conversations through the switch at any given time. Hubs and bridges offer only limited numbers of simultaneous conversations at a time.

Remember that if two stations try to communicate in a hubbed environment, *collisions* result. In a switched environment, however, collisions don't occur because each workstation has its own dedicated path through the network. The network actually has dramatically more bandwidth, and each station now has a dedicated 10-Mbps worth of bandwidth.

4.1.8 Switches vs. Hubs

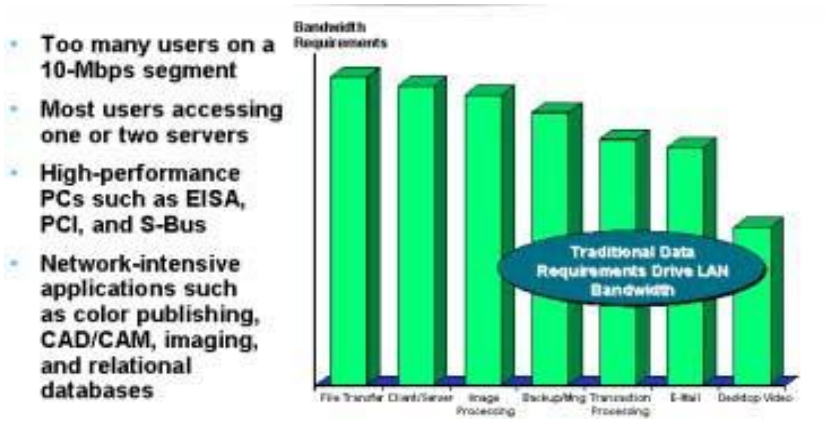
Figure 1: Switches vs. Hubs



Remember when considering switches versus hubs that with hubs all traffic fights for the same fixed amount of bandwidth. Figure [1] shows improved traffic flow through the network because each workstation has a dedicated lane.

4.1.9 Typical Causes of Network Congestion

Figure 1: Typical Causes of Network Congestion



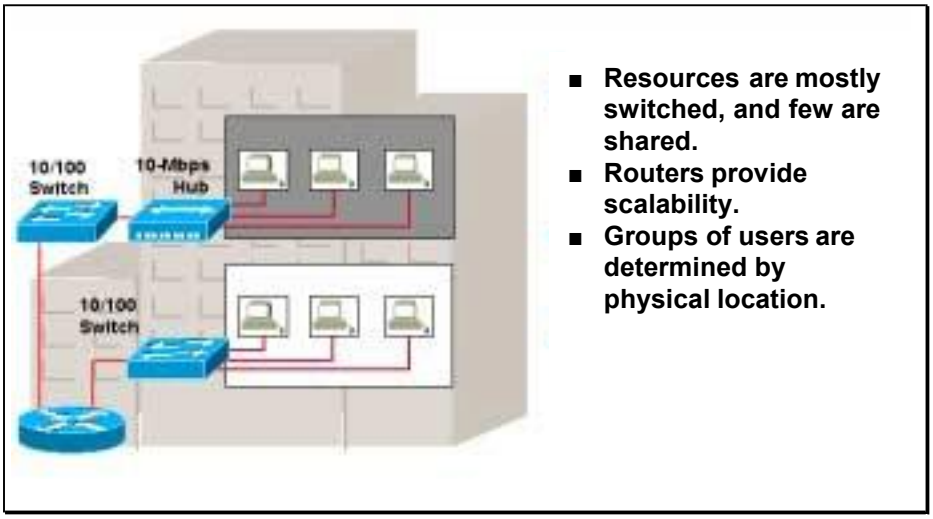
How do congestion problems manifest themselves in a network? Remember that shared LAN segments have a fixed amount of bandwidth. As users are added, proportionally, the amount of bandwidth per user decreases, and the result is collisions—and, of course, collisions reduce performance.

Now consider the newer technologies used in workstations. With early LAN technologies, workstations were relatively limited in terms of the amount of traffic they could deliver to the network. But with newer, faster CPUs, faster buses, faster peripherals, and so on, it is much easier for a single workstation to fill up a network segment. So with faster PCs, applications can be used to better advantage—at the expense of reduced available bandwidth.

In particular, bandwidth-intensive applications that are used today, such as desktop publishing, engineering applications, imaging applications, and even multimedia applications, deplete available bandwidth faster than ever.

4.1.10 Today's LANs

Figure 1: Today's LANs



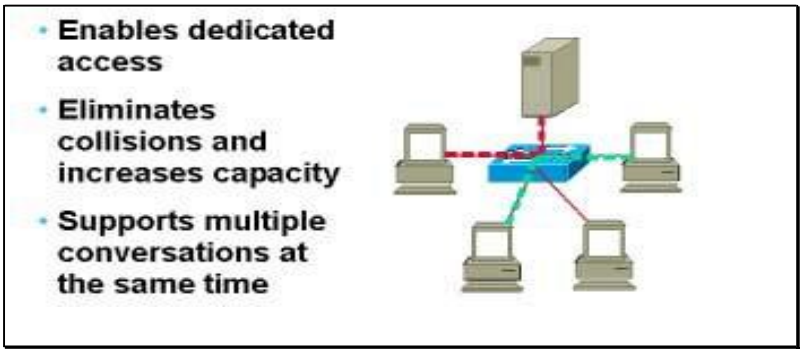
Switched infrastructures are the most commonly implemented LANs today. Because of the price point of deploying switches, many companies are bypassing the shared-hub technologies and moving directly to switches. Even within switched networks, at some point routers are needed to provide scalability. In addition, grouping of users is largely determined by physical location.

Thus we have seen the limitations of traditional shared LAN technologies. Now let's see how we can improve performance in some of these areas. Consider deployment of LAN switches to take advantage of some new, improved technologies.

4.2 LAN Switching Basics

4.2.1 Microsegmentation

Figure 1: LAN Switching Basics



Again, LAN switching provides *microsegmentation*, which gives dedicated bandwidth to each user on the network. Microsegmentation eliminates collisions in a network, and effectively increases the capacity for each station connected to the network. It also supports multiple, simultaneous conversations at any given time, resulting in dramatic improvement in available bandwidth and scalability.

4.2.2 LAN Switch Operation

Figure 1: LAN Switch Operation

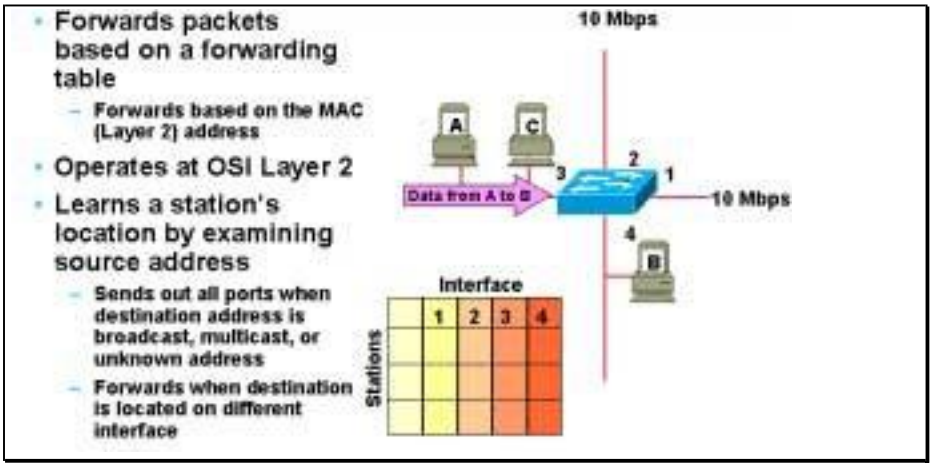


Figure 2: LAN Switch Operation

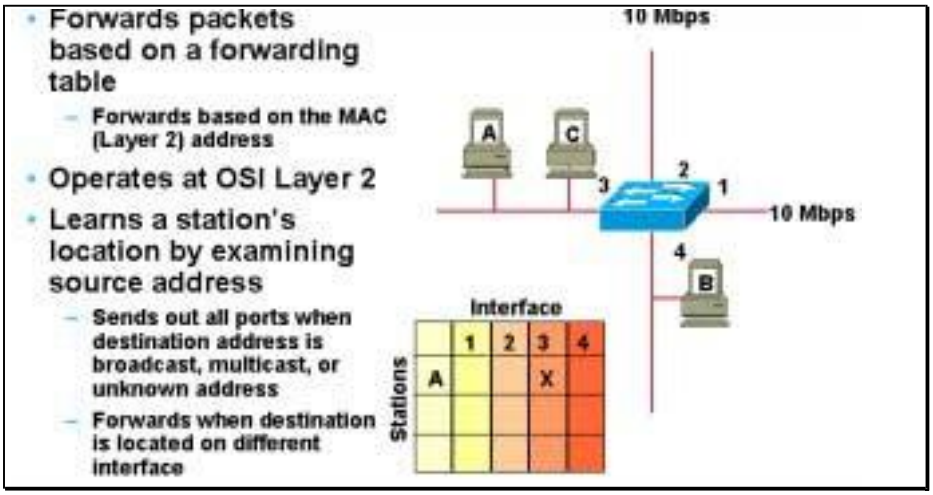


Figure 3: LAN Switch Operation

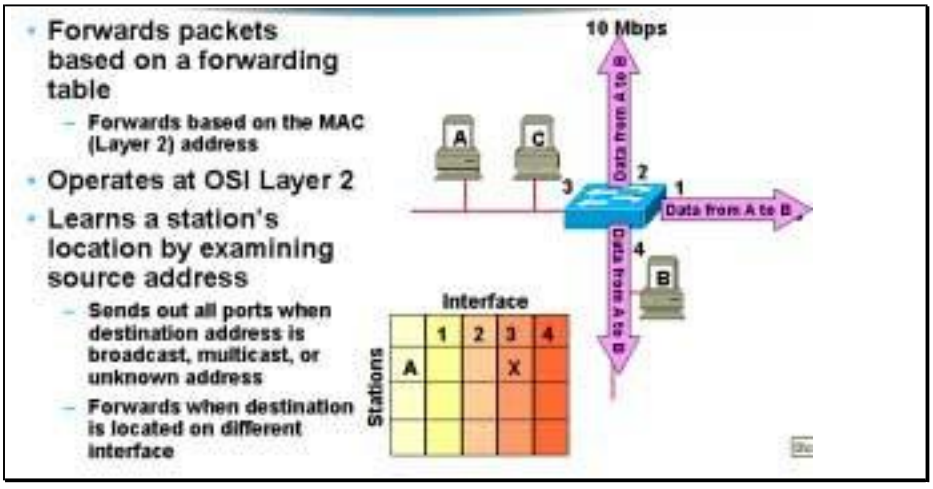


Figure 4: LAN Switch Operation

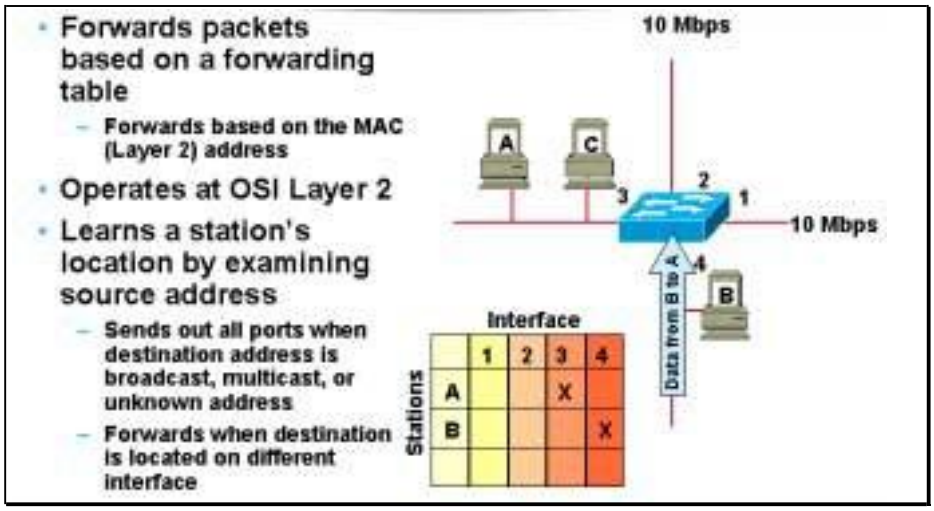
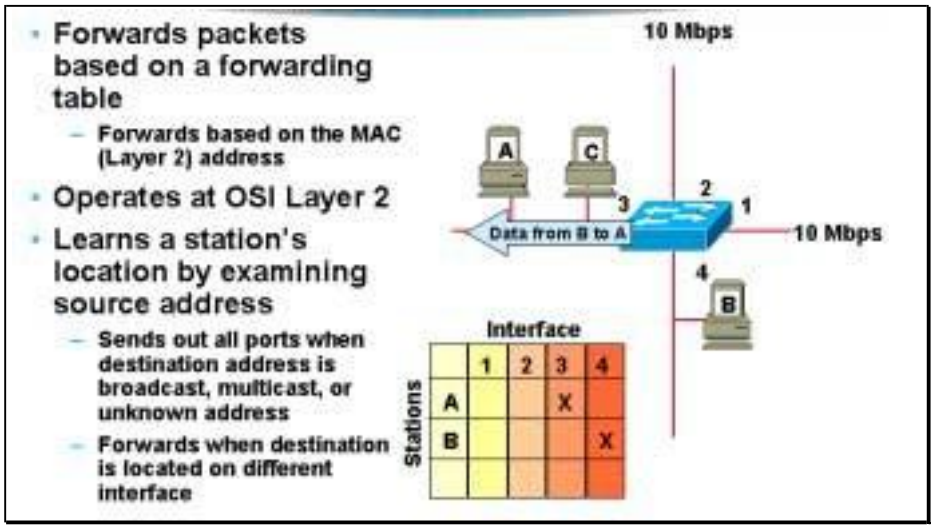


Figure 5: LAN Switch Operation



Now let's look at the fundamental operation of a LAN switch. As indicated in Figure **[1]**, some data needs to be transmitted from Station A to Station B. Remember that as this traffic goes through the network, the switch operates at Layer 2, meaning that the switch can look at the Media Access Control (MAC)-layer address. The switch actually looks at the traffic as it goes through to discover the MAC address and store it in an address table (see Figure **[2]**).

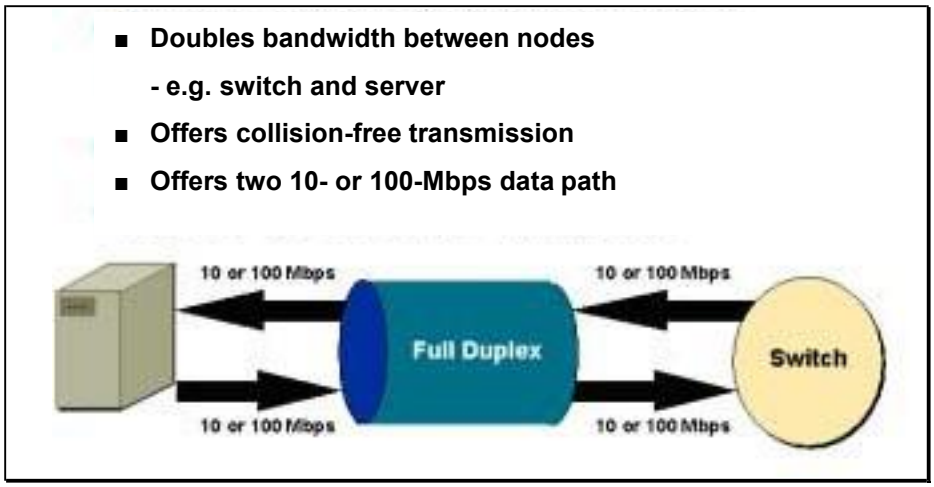
So, as the traffic goes through, an entry is made in this table in terms of station and the port that it's connected to on the switch.

When that frame of data is in the switch, it floods to all ports because the destination station is unknown. After the address entry is made in the table, however, a response comes back from Station B to Station A, and now the switch knows where Station A is connected to the network (see Figure **[3]**).

So the data is transmitted into the switch, but notice that the switch doesn't flood the traffic this time—it sends the data out only port 3, because it knows where Station A is on the network (see Figures **[4]** and **[5]**). The original transmission indicated where that MAC address came from, allowing the switch to more efficiently deliver traffic in the network. As noted previously, the fundamental concept behind LAN switching is called *microsegmentation*, and it allows multiple, simultaneous conversations in the network.

4.2.3 Switching Technology: Full Duplex

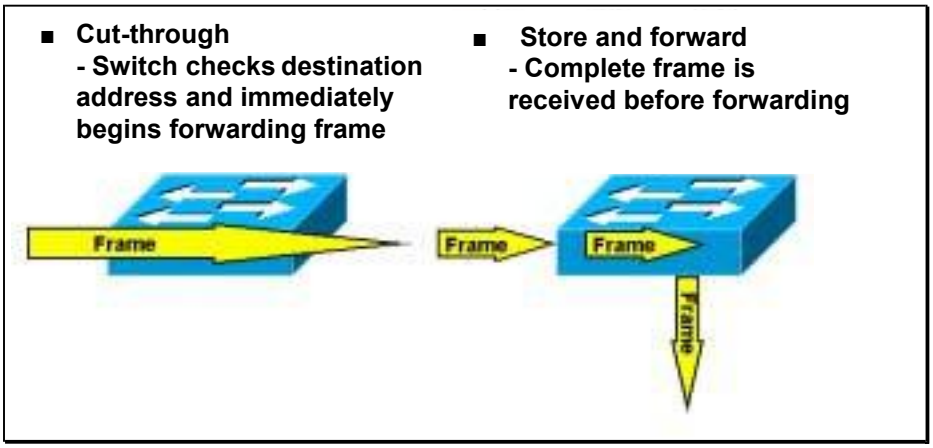
Figure 1: Switching Technology: Full Duplex



Another concept of LAN switching that dramatically improves scalability is *full-duplex transmission*, which effectively doubles the amount of bandwidth between nodes. This feature can be important, for example, between high-bandwidth consumers such as between a switch and a server connection. It provides essentially collision-free transmissions in the network. In 10-Mbps connections, for example, it effectively provides 10 Mb of transmit capacity and 10 Mb of receive capacity, for effectively 20 Mb of capacity on a single connection. Likewise, a 100-Mbps connection offers effectively 200 Mbps of throughput.

4.2.4 Switching Technology: Two Methods

Figure 1: Switching Technology: Two Methods



There are two modes of switching, offering different performance and latency.(*Note*: Latency, sometimes called propagation delay, is the time a frame, or packet,of data takes to travel from the source station or node to its final destination on the network.)

First, in *cut-through switching*, the switch reads the destination MAC address as the traffic flows through the switch and "cuts through" to its destination without continuing to read the rest of the frame. Cut-through switching offers better performance than the second method, known as *store and forward*.

In store-and-forward switching, the switch reads the entire frame of data, decides where it needs to go, and sends it on its way. The obvious trade-off here is the longer time it takes the switch to read the entire frame. As it reads the entire frame, however, it performs some error correction on that frame, possibly increasing reliability. In summary, although cut-through switching is faster, it offers no error detection.

4.2.5 The Need for Spanning Tree

Figure 1: The Need for Spanning Tree

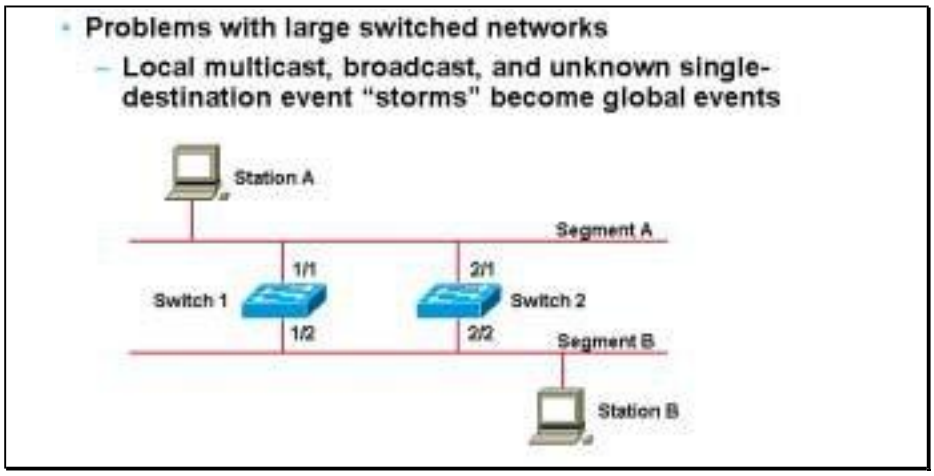
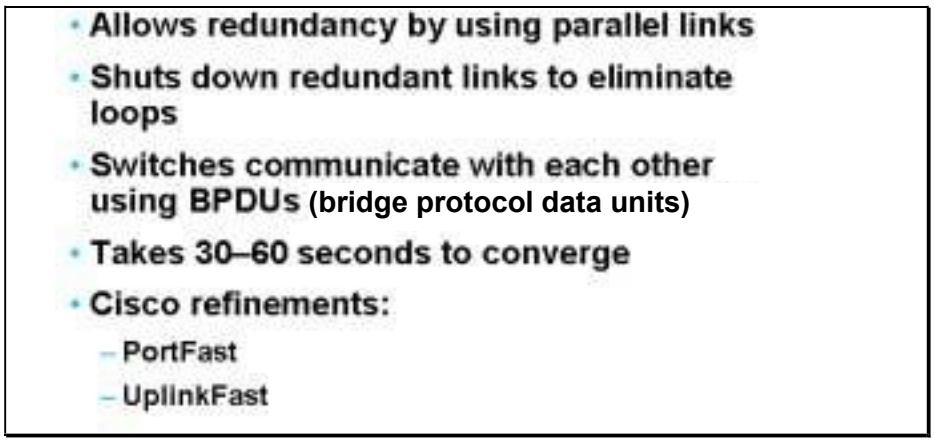


Figure 2: 802.1d Spanning-Tree Protocol (STP)



Now let's look at some key technologies within LAN switching. In large networks, one of the problems at Layer 2 in the OSI model is that if forwarding decisions are made only at this layer, the network cannot have any physical layer loops.

Thus in a simple network, as we see in Figure **[1]**, when a switch has any multicast, broadcast, or any unknown traffic, the result will be storms of traffic being looped endlessly through the network. To prevent this situation, loops need to be eliminated. One way to eliminate the loops would be to physically disconnect those segments, but that is obviously not a good solution because there would be no physical redundancy in the network. Thus what is needed is a way to *logically* cut out the loops in the network so that they can be reenabled *dynamically* if necessary.

802.1d Spanning-Tree
The solution is the Spanning-Tree Protocol, or STP. STP is actually an industry standard defined by the IEEE standards committee, known as the 802.1d Spanning-Tree Protocol. STP allows physical redundancy in the network, but it logically disconnects those loops.

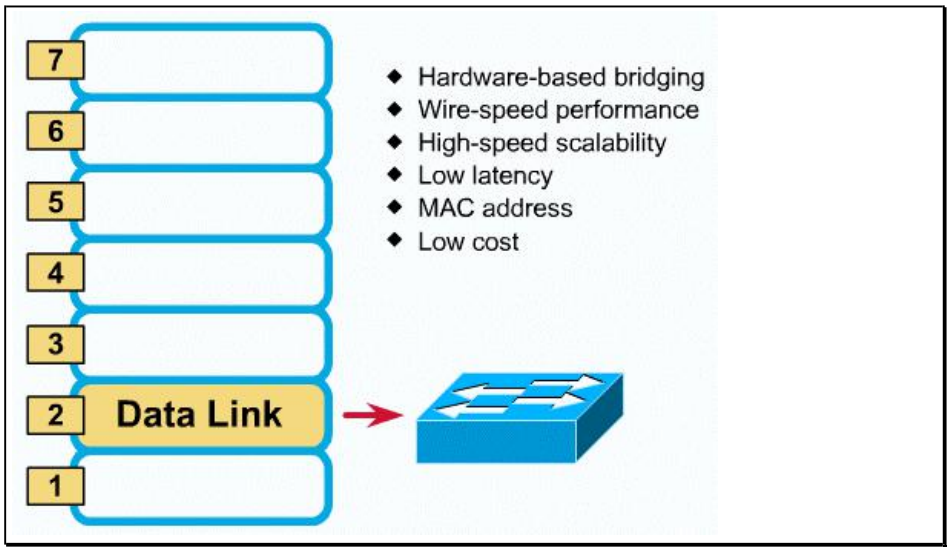
It's important to understand that logically disconnecting the loops allows dynamic reestablishment of a connection if a failure occurs within the network. Switches and bridges can disconnect loops simply by communicating back and forth with *hello* messages. Hello messages are status messages that the bridges and switches exchange periodically so they know the status of those logical connections and disconnections. If a switch or bridge stops hearing a given communication from a certain device on the network, that network device has failed. And when a network failure occurs, the link must be reestablished in order to maintain redundancy. Technically, these little exchanges are known as bridge protocol data units, or BPDUs.

Although STP works well, it can take from 30 seconds to a full minute for the network to fully converge—in order for all devices to know the status of the network.

4.3 Multilayer Switching Devices

4.3.1 Layer 2 Switching Devices

Figure 1: Layer 2 Switching



A Layer 2 switch is operationally similar to a multiport bridge, but has a much higher capacity and supports many new features, such as full-duplex operation. A Layer 2 LAN switch performs switching and filtering based on the OSI data link layer (Layer 2) MAC address. Like bridges, Layer 2 switches are completely transparent to network protocols and user applications.

Bridges and switches analyze incoming frames, make forwarding decisions based on information contained in the frames, and forward the frames toward the destination. Upper-layer protocol transparency is a primary advantage of both bridging and switching. Because both device types operate at the data link layer, they are not required to examine upper-layer information. Bridges are also capable of filtering frames based on any Layer 2 fields.

Although bridges and switches share most relevant attributes, several distinctions differentiate these technologies. Switches are significantly faster because they switch in hardware. Switches also can support higher port densities than bridges. And, some switches support cut-through switching, reducing latency and delays in the network, whereas bridges support only store-and-forward traffic switching. The primary differences, though, are that bridges perform switching via software (as opposed to hardware) and switches have a higher port density.

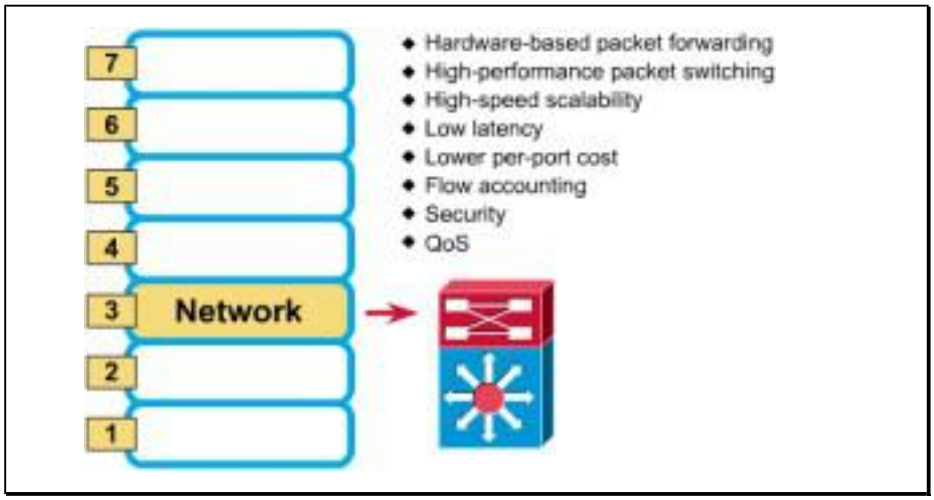
Layer 2 switching is basically hardware-based bridging. In a switch, frame forwarding is handled by specialized hardware called application-specific integrated circuits (ASICs). The ASIC technology engineered for switches allows for scalability up to gigabit speeds, with low latency at costs significantly lower than Ethernet bridges.

Layer 2 switches provide network managers with the ability to increase bandwidth without adding complexity to the network. Layer 2 data frames consist of both control information, such as MAC addresses, and end-user content. At Layer 2, no modification of the frame control information is required when moving between similar Layer 1 interfaces, such as Ethernet and Fast Ethernet. However, changes to control information may occur when bridging between unlike LAN types such as Fiber Distributed Data Interface (FDDI) or ATM and Ethernet.

Workgroup connectivity and network segmentation are the two primary uses for Layer 2 switches. The high performance of a Layer 2 switch allows for network designs that significantly decrease the number of hosts per physical segment. Decreasing the number of hosts per segment leads to a flatter design with more segments in the campus network. However, despite the advantages of Layer 2 switching, it still has all the same characteristics and limitations of legacy bridging.

4.3.2 Layer 3 Switching Devices

Figure 1: Layer 3 Switching



Layer 3 switches are, essentially, a cross between a LAN switch and a router. Each port on the switch is a separate LAN port, but the forwarding engine actually calculates and stores routes based on IP addresses, not MAC addresses. You can think of a Layer 3 switch as a switch that also performs hardware-based routing using Layer 3 (network) addresses.

Layer 3 switches available today tend to support only IP or both IP and Internetwork Packet Exchange (IPX), to the exclusion of other network layer protocols. Similarly, selection of LAN port technologies is frequently limited to 10-, 100-, or 1000- Mbps Ethernet.

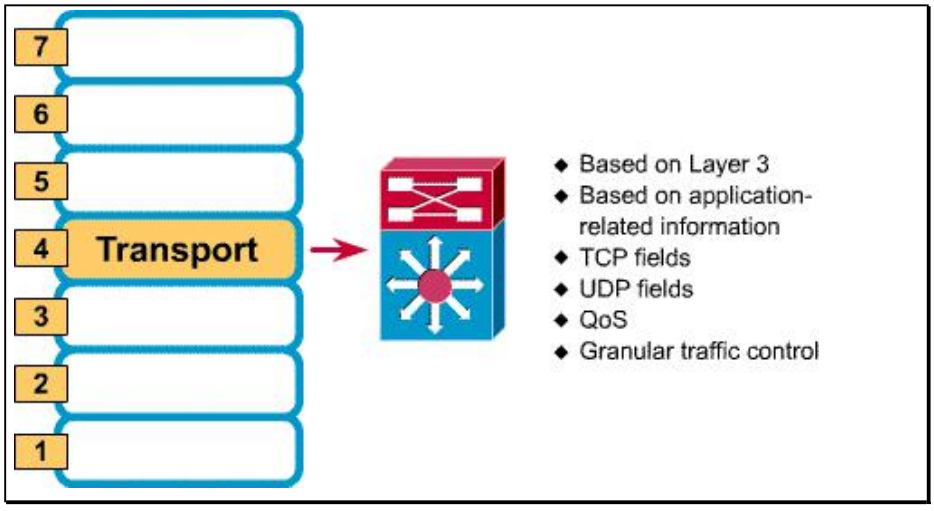
Basically, Layer 3 switching is hardware-based routing. In particular, the packet forwarding is handled by specialized hardware ASICs. The goal is to capture the speed of switching and the scalability of routing. A Layer 3 switch acts on a packet in the same way that a traditional router does; for example:

- Determining the forwarding path based on Layer 3 information
- Validating the integrity of the Layer 3 header via checksum
- Verifying packet expiration and updates accordingly
- Processing and responding to any option information
- Updating forwarding statistics in the Management Information Base (MIB)
- Applying security controls if required
- Implementing quality of service (QoS)

The primary difference between the packet-switching operation of a router and a Layer 3 switch lies in the physical implementation. In general-purpose routers, microprocessor-based engines typically perform software-based packet switching. A Layer 3 switch performs packet switching with hardware. Because it is designed to handle high-performance LAN traffic, a Layer 3 switch can be placed anywhere within the network, offering a cost-effective alternative to the traditional router.

4.3.3 Layer 4 Switching Devices

Figure 1: Layer 4 Switching



Layer 4 switching refers to Layer 3 hardware-based routing that accounts for Layer 4 control information. Information in packet headers typically includes Layer 3 addressing, the Layer 3 protocol type, and more fields relevant to Layer 3 devices, such as Time To Live (TTL) and checksum. The packet also contains information relevant to the higher layers within the communicating hosts, such as the protocol type and port number.

A simple definition of Layer 4 switching is the ability to make forwarding decisions based not just on the MAC address or source/destination IP addresses, but on Layer 4 parameters such as port numbers as well. In TCP or User Datagram Protocol (UDP) flows, the application is encoded as a port number in the TCP or UDP header.

Routers are capable of controlling traffic based on Layer 4 information. One method of controlling Layer 4 traffic is by using extended access lists. Another method of providing Layer 4 accounting of flows is available, NetFlow Switching, which is utilized on the Cisco 7200 and 7500 Router platforms.

Finally, when performing Layer 4 functions, a switch reads the TCP and UDP fields within the headers to determine what type of information the packet is carrying. The network manager can program the switch to prioritize traffic by application. This function allows network managers to define a quality of service (QoS) for end users. When used for QoS purposes, Layer 4 switching might mean that a videoconferencing application is granted more bandwidth than an e-mail message or File Transfer Protocol (FTP) packet.

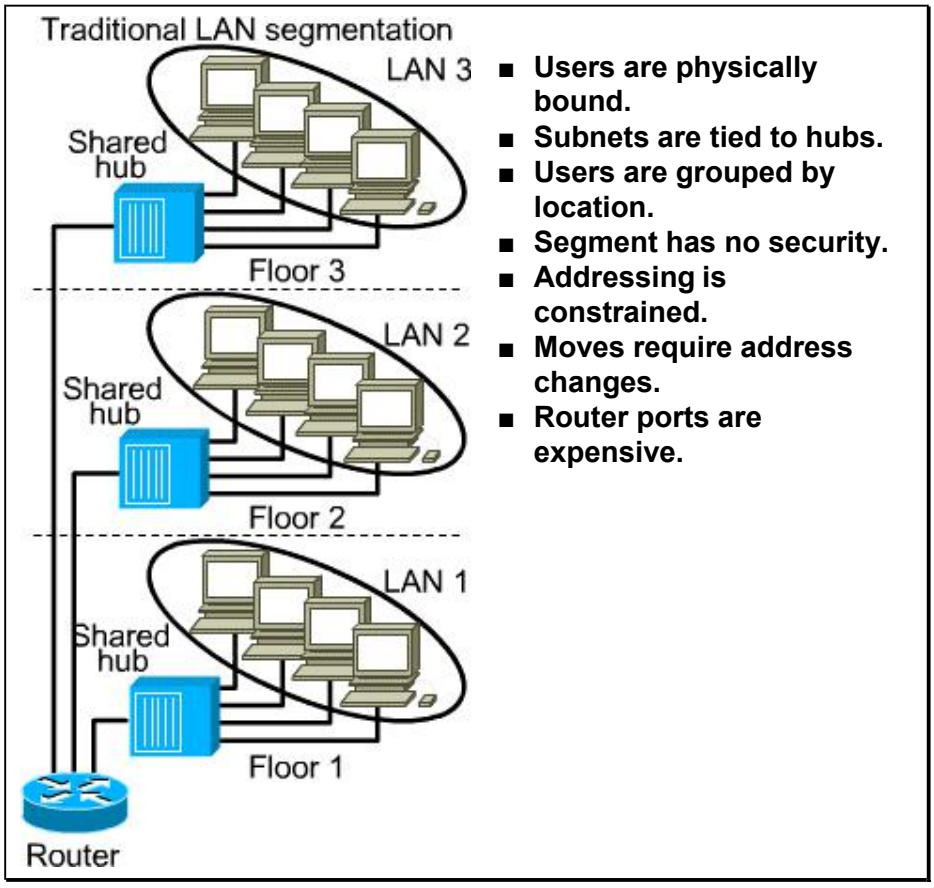
Layer 4 switching is necessary if your policy dictates granular control of traffic by application or if you require accounting of traffic itemized in terms of applications. However, it should be noted that switches performing Layer 4 switching need the ability to identify and store large numbers of forwarding-table entries, especially if the switch is within the core of an enterprise network. Many Layer 2 and Layer 3 switches have forwarding tables that are sized in proportion to the number of network devices.

With Layer 4 switches, the number of network devices must be multiplied by the number of different application protocols and conversations in use in the network. Thus, the size of the forwarding table can grow quickly as the numbers of end devices and types of applications increase. This large table capacity is essential to creating a high-performance switch that supports wire-speed Layer 4 forwarding of traffic.

4.4 Virtual LANs

4.4.1 Constraints of Shared LANs

Figure 1: Constraints of Shared LANs



Let's begin by reviewing some of the limitations of traditional, shared local-area networks. Users are generally bound by their physical location in a network; that is, the actual port or hub that they plug into determines what resources they can connect to and how they're grouped together in a LAN. Also, users are generally grouped not logically, but physically by where they sit and where they gain their physical connectivity.

Shared LAN networks also offer very little security, because on a hub or concentrator all traffic in the network is available on all ports. That's the inherent nature of shared LAN devices. Also, there are constraints with addressing because of the physical layout and requirements in the shared technology environment, and moves, adds, and changes can be very difficult as well because they require making changes either on a patch panel or in a wiring closet—wherever the hubs or concentrators reside.

Lastly, routers are needed to connect different segments together. So if separation occurs, router ports either may not be available, or if they are, they're relatively expensive compared to some of the alternatives. These are some of the reasons that virtual LANs, or VLANs, are implemented.

4.4.2 Virtual LANs

Figure 1: Virtual LANs

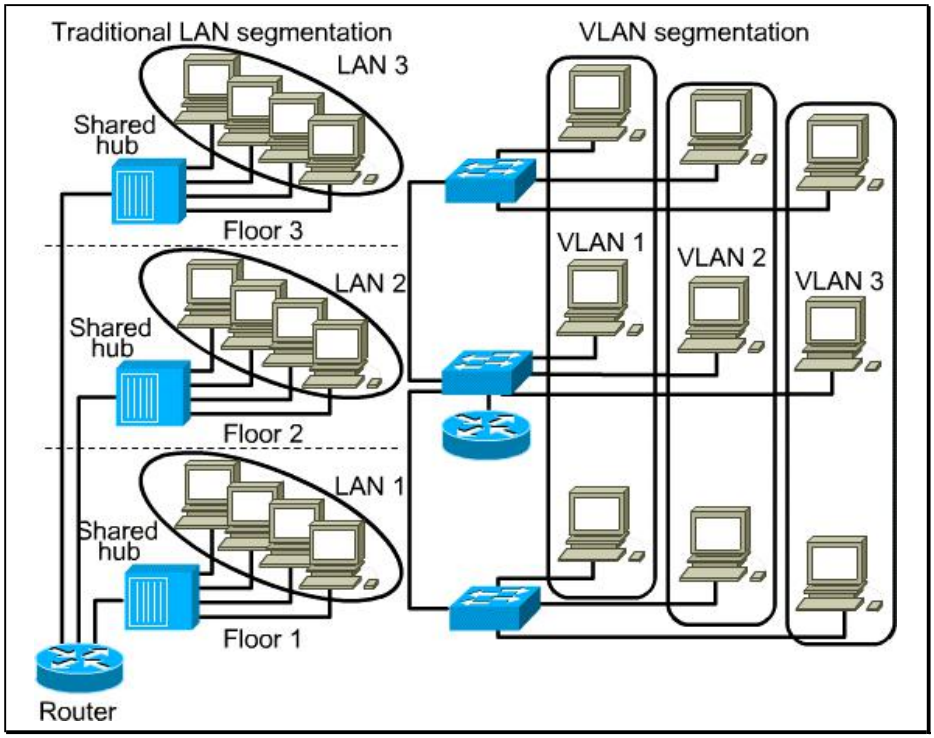
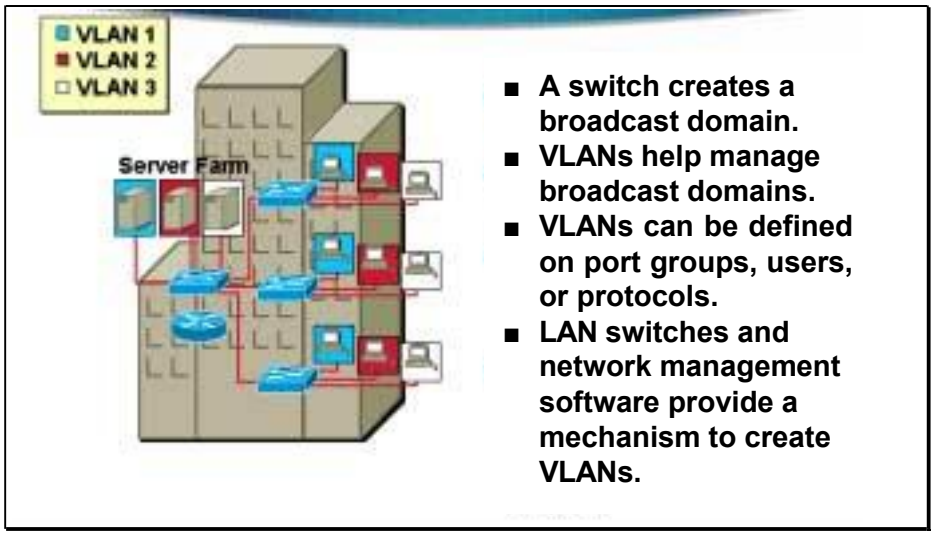


Figure 1: Virtual LANs



A Virtual LAN (VLAN) is a logical grouping of devices or users, as shown in Figure [1]. These devices or users can be grouped by function, department, or application, regardless of their physical segment location. VLAN configuration is done at the switch via software. VLANs are not standardized and require the use of proprietary software from the switch vendor.

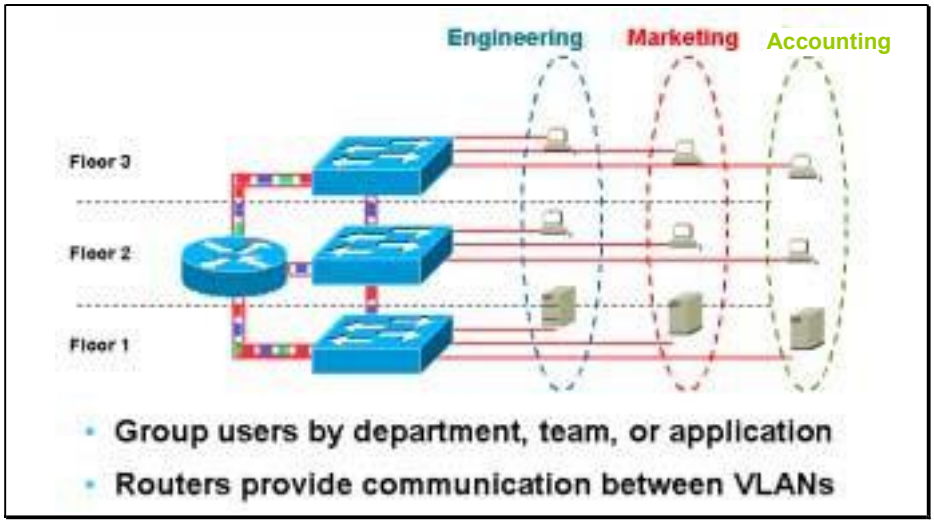
As mentioned in previous section, a typical shared LAN is configured according to the physical infrastructure it is connecting. Users are grouped based on their

location in relation to the hub they are plugged in to and how the cable is run to the wiring closet. The router interconnecting each shared hub typically provides segmentation and can act as a broadcast firewall. The segments created by switches do not. Traditional LAN segmentation does not group users according to their workgroup association or need for bandwidth. Therefore, they share the same segment and contend for the same bandwidth, although the bandwidth requirements may vary greatly by workgroup or department.

VLANs take a single broadcast domain and limit it within a given switch so that multiple segments can exist within a switch instead of one device providing a single broadcast domain. VLANs also can help manage broadcast traffic. Because the switch itself without VLANs propagates all broadcast traffic, as soon as multiple VLANs are created, they will block the propagation of that broadcast traffic. Thus VLANs can help contain broadcast traffic. In addition, VLANs can be defined by port groups, by actual user IDs, by MAC address, or even by protocol. VLAN group membership can be defined in several ways. Lastly, LAN switches and network management software provide a mechanism to actually create, and more importantly, to manage VLANs over the long term (see Figure **[2]**).

4.4.3 Remove the Physical Boundaries

Figure 1: Remove the Physical Boundaries



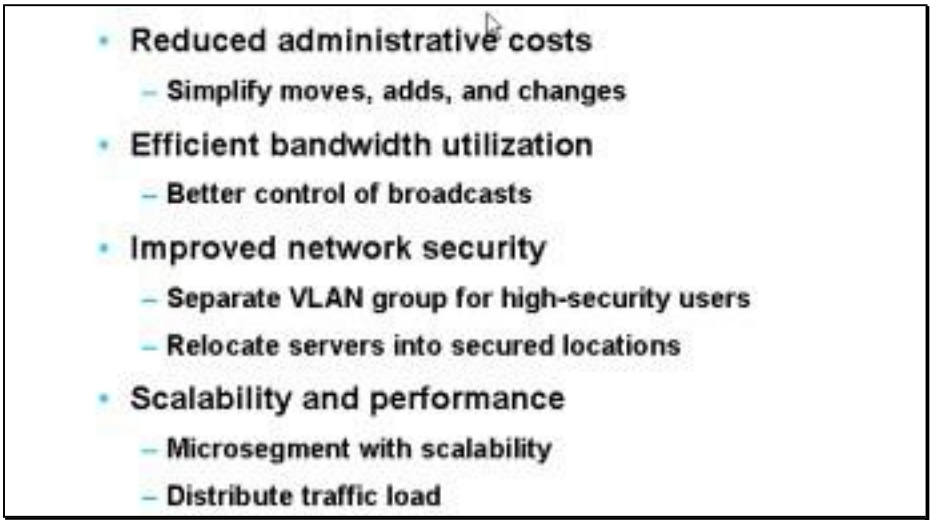
Conceptually, VLANs provide greater segmentation and organizational flexibility. VLAN technology allows you to group switch ports and the users connected to them into logically defined communities of interest. These groupings can be coworkers within the same department, a cross-functional product team, or diverse users sharing the same network application or software.

Grouping these ports and users into communities of interest—referred to as VLAN organizations—can be accomplished within a single switch, or more powerfully, between connected switches within the enterprise. By grouping ports and users together across multiple switches, VLANs can span single building infrastructures or interconnected buildings. As shown in Figure **[1]**, VLANs completely remove the physical constraints of workgroup communications across the enterprise.

Additionally, the role of the router evolves beyond the more traditional role of firewalls and broadcast suppression to policy-based control, broadcast management, and route processing and distribution. Equally as important, routers remain vital for switched architectures configured as VLANs because they provide the communication between logically defined workgroups (VLANs). Routers also provide VLAN access to shared resources such as servers and hosts, and connect to other parts of the network that are either logically segmented with the more traditional subnet approach or require access to remote sites across wide-area links. Layer 3 communication, either embedded in the switch or provided externally, is an integral part of any high-performance switching architecture.

4.4.4 VLAN Benefits

Figure 1: VLAN Benefits

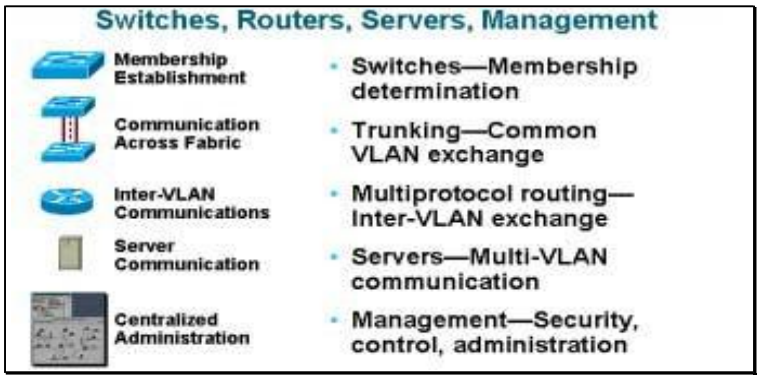


Some of the key benefits of using VLANs are given in Figure [1]. The original motivation for VLANs was to reduce the administrative costs associated with managing a routine shared network. What was needed was a way to simplify the moves, adds, and changes that were commonly associated with most organizations as their networks evolved. VLANs offer other benefits, including better bandwidth control. Segmenting a switch and into multiple VLANs limits the size of broadcast domains. In other words, it limits how far and to how many ports the broadcast traffic is propagated.

Another benefit of VLANs is improved network security. VLANs can be separated on the switch, so that traffic from one VLAN is not communicated to another VLAN. In addition, servers can be relocated into secured locations and connectivity provided to only those workstations that need it. VLANs can improve scalability and performance, and microsegmentation can dramatically improve some key performance aspects in a LAN. Finally, VLANs can be used to distribute the traffic load more efficiently throughout the LAN.

4.4.5 VLAN Components

Figure 1: VLAN Components



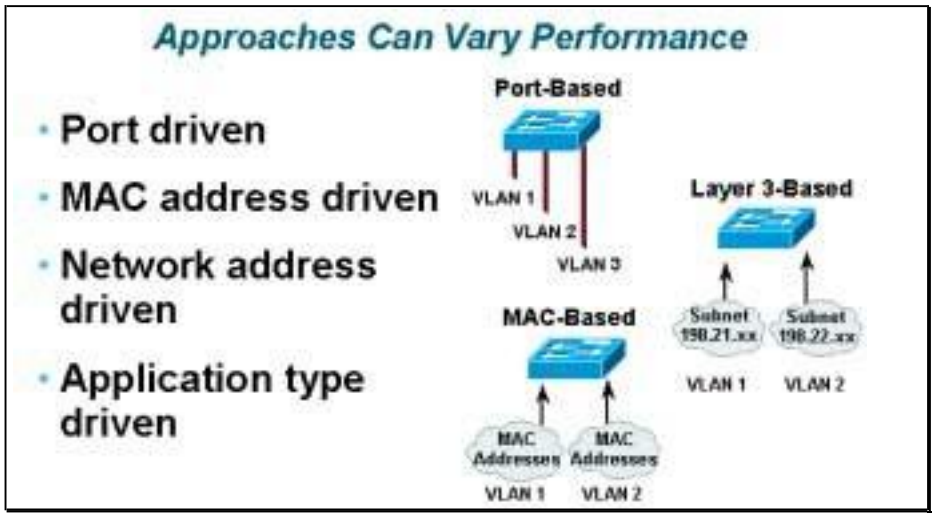
To truly understand VLANs, we need to understand some of the key components. Switches play a role, but also routers, servers, and management stations play a role in successfully deploying VLANs.

Switches determine the VLAN membership and provide the basic connectivity of the VLAN members. Trunking functionality is also needed. In other words, a way is needed to exchange VLAN information between switches if the network VLANs span multiple physical switches. Also, VLANs require multiprotocol routing functionality. The routers provide this functionality.

Remember that VLANs essentially create multiple network segments, so a method is needed to route traffic between the VLANs, and that's what the router does. Traffic flow in the network can be optimized by giving servers the ability to discriminate traffic down to the individual VLAN level. Lastly, network management functionality is needed in order to initially deploy and then manage VLANs. Thus it's not just the switch that enables a successful VLAN deployment.

4.4.6 Establishing VLAN Membership

Figure 1: Establishing VLAN Membership



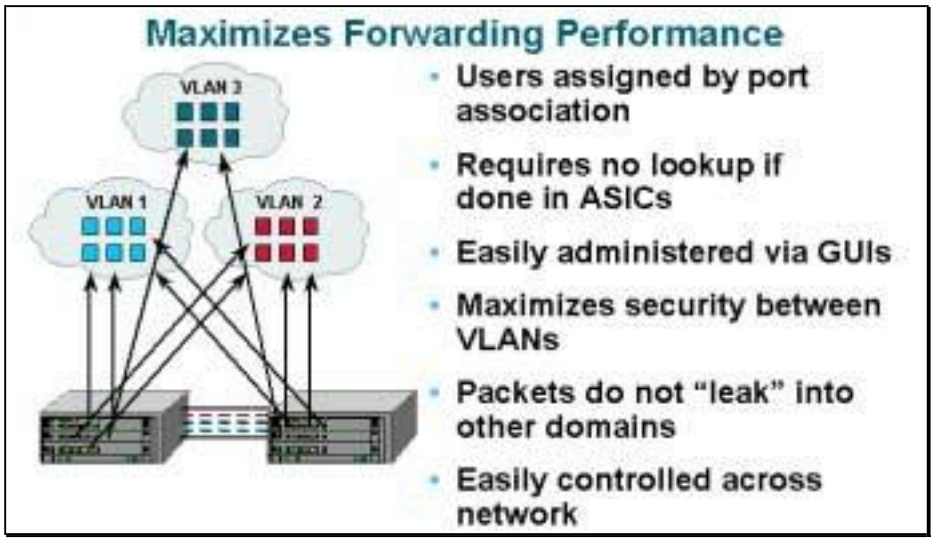
VLAN membership can be established in several ways, all involving trade-offs. In port-driven VLAN membership, VLANs are determined by the port that a given workstation plugs into.

For example, we may say on a 12-port switch that ports 1 through 6 are VLAN 1, and ports 7 through 12 are VLAN 2.

VLAN membership can also be defined by MAC address. That is, the switch looks at a MAC address and then dynamically determines which VLAN a station belongs to based on its MAC address. This scenario offers a mechanism for dynamic VLAN membership, similar to network address VLAN membership. We can look at a workstation IP address, for example, or a user ID as the user logs into the network to determine which VLAN a station belongs to. Lastly, we can even look at the application type.

4.4.7 Membership by Port

Figure 1: Membership by Port

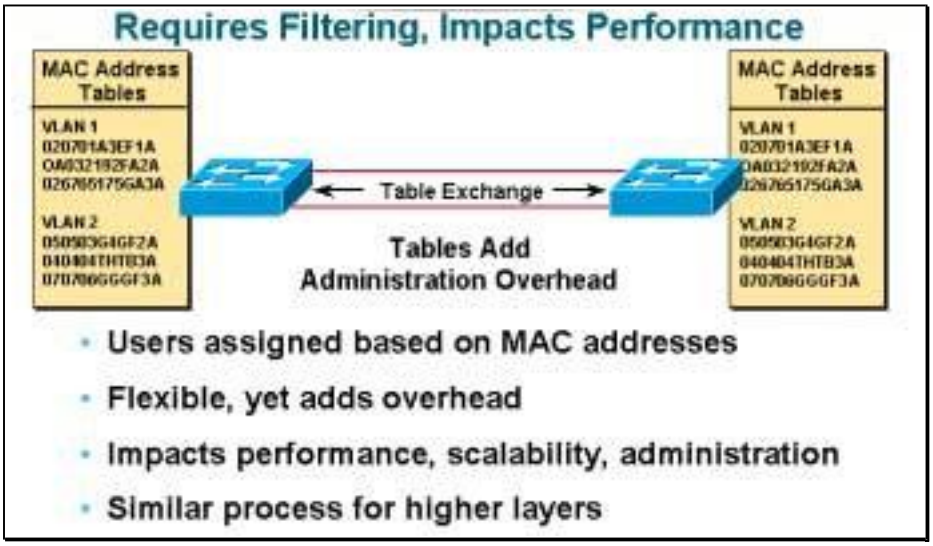


Providing the maximum forwarding performance, membership by port is the simplest mechanism of VLAN port membership. Users are simply assigned by the port that they plug into. No address lookups are required in the ASICs because the administrator manually defines the VLAN a particular port belongs to. Port-based VLAN member is also known as static VLANs.

Administration is relatively easy, done either by command-line interface or by a graphical user interface (GUI). Port membership can also maximize the security between VLANs, and port-based VLANs can be created to ensure that the packets don't leak into other domains. Network administration is easily controlled across the entire network with port membership.

4.4.8 Membership by MAC Addresses

Figure 1: Membership by MAC Addresses

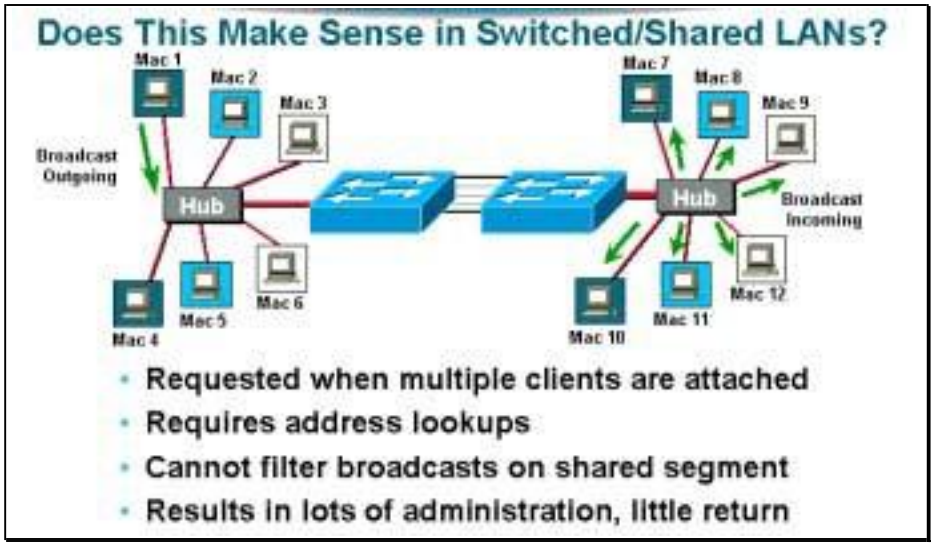


VLAN membership can also be determined by MAC address. This scenario requires filtering because we have to look at the traffic as it goes through the switch, a process that impacts performance. In addition, we have to look up in a table to determine which VLAN that traffic belongs to.

Although this process offers flexibility, it also adds to the switch processing overhead. It offers flexibility to support mobile users and dynamically determine their VLAN membership based on the port that they plug into. The trade-offs lie in the areas of performance, scalability, administration, and so on.

4.4.9 Multiple VLANs per Port

Figure 1: Multiple VLANs per Port



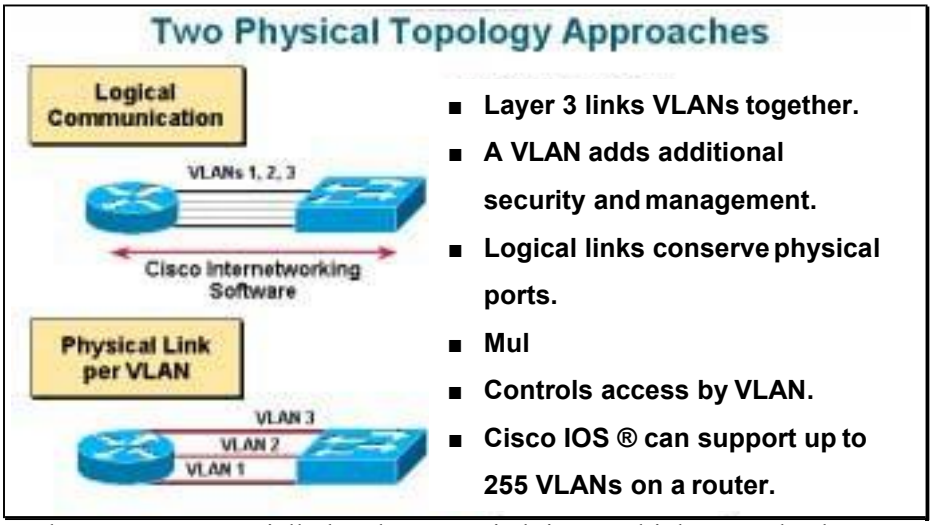
Multiple VLANs can be defined on a given port. Although this solution might be required in situations where hubs are plugged into switch ports, it might defeat the original purpose of creating the VLAN.

Notice that several VLANs are defined with individual VLAN members connected to the hubs, not to the switches. Now consider the broadcast message that will be sent by MAC station 1 on the left side of Figure **[1]**.

As that broadcast message is propagated through the network, when it reaches the hub on the other side of the network, that hub will flood that broadcast traffic to all stations on that hub. Thus defining multiple VLANs on a single port negates the purpose of limiting broadcast traffic—and a key advantage of having VLANs is lost. Although this setup may be required in some instances for connectivity, remember that its use might actually defeat one of the key purposes of a VLAN.

4.4.10 Communicating Between VLANs

Figure 1: Communicating Between VLANs



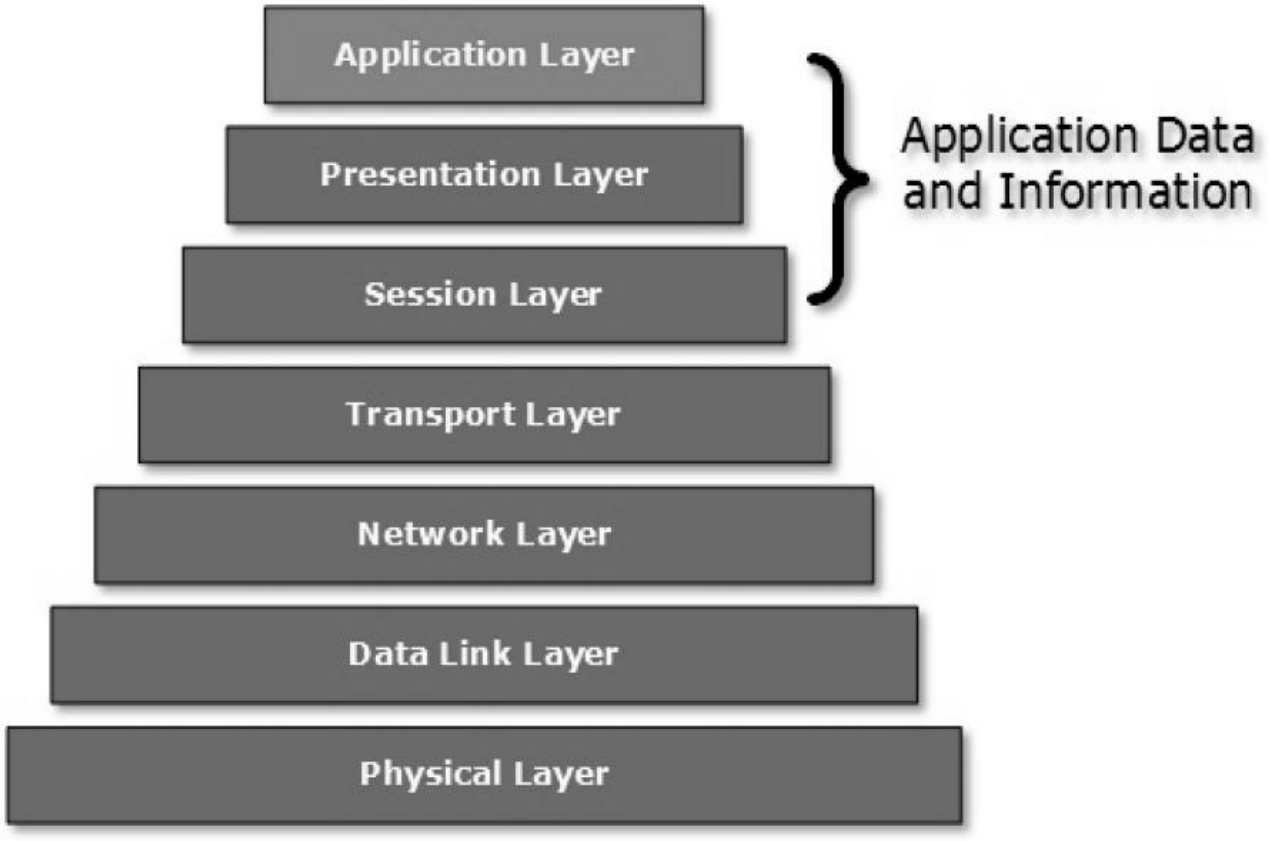
Remember that VLANs essentially break up a switch into multiple, completely separate network segments. Now a router is needed to provide connectivity between those network segments.

Connectivity can be achieved in two ways, logical connectivity or physical connectivity. Logical connectivity involves a single connection to the router, and that single connection can support multiple VLANs. This configuration is sometimes referred to as a “router on a stick,” because there is a single connection to the router but multiple logical connections inside that physical connection. In addition, the router performs the routing among the multiple VLANs.

Another connectivity option is to have a separate physical connection for each VLAN. The obvious trade-off is that this configuration requires a separate physical port for each of the VLANs. Although this setup may provide better separation and better performance, the trade-off is the requirement for more resources, namely the individual interfaces on the router. With regard to VLAN support on the router, Cisco IOS® Software, for example, can support up to 255 VLANs on a given router. Thus there is flexibility in terms of routing between VLANs.

APPLICATION LAYER

Application layer is the topmost layer in OSI and TCP/IP layered model. This layer exists in both layered Models because of its significance, of interacting with user and user applications. This layer is for applications which are involved in communication system.
A user may or may not directly interacts with the applications. Application layer is where the actual communication is initiated and reflects. Because this layer is on the top of the layer stack, it does not serve any other layers. Application layer takes the help of Transport and all layers below it to communicate or transfer its data to the remote host.
When an application layer protocol wants to communicate with its peer application layer protocol on remote host, it hands over the data or information to the Transport layer. The transport layer does the rest with the help of all the layers below it.



There is an ambiguity in understanding Application Layer and its protocol. Not every user application can be put into Application Layer except those applications which interact with the communication system. For example, designing software or text-editor cannot be considered as application layer programs.

On the other hand, when we use a Web Browser, which is actually using Hyper Text Transfer Protocol (HTTP) to interact with the network, HTTP is Application Layer protocol.

Another example is File Transfer Protocol, which helps a user to transfer text based or binary files across the network. A user can use this protocol in either GUI based software like FileZilla or CuteFTP and the same user can use FTP in Command Line mode.
Hence, irrespective of which software you use, it is the protocol which is considered at Application Layer used by that software. DNS is a protocol which helps user application protocols such as HTTP to accomplish its work.

APPLICATION PROTOCOLS

There are several protocols which work for users in Application Layer. Application layer protocols can be broadly divided into two categories:

Protocols which are used by users. For example, eMail.

Protocols which help and support protocols used by users. For example, DNS.

Few of Application layer protocols are described below:

Domain Name System

The Domain Name System (DNS) works on Client Server model. It uses UDP protocol for transport layer communication. DNS uses hierarchical domain based naming scheme. The DNS server is configured with Fully Qualified Domain Names (FQDN) and email addresses mapped with their respective Internet Protocol addresses.
A DNS server is requested with FQDN and it responds back with the IP address mapped with it. DNS uses UDP port 53.

Simple Mail Transfer Protocol

The Simple Mail Transfer Protocol (SMTP) is used to transfer electronic mail from one user to another. This task is done by means of email client software (User Agents) the user is using. User Agents help the user to type and format the email and store it until internet is available. When an email is submitted to send, the sending process is handled by Message Transfer Agent which is normally comes inbuilt in email client software.
Message Transfer Agent uses SMTP to forward the email to another Message Transfer Agent (Server side). While SMTP is used by end user to only send the emails, the Servers normally use SMTP to send as well as receive emails. SMTP uses TCP port number 25 and 587.
Client software uses Internet Message Access Protocol (IMAP) or POP protocols to receive emails.

File Transfer Protocol

The File Transfer Protocol (FTP) is the most widely used protocol for file transfer over the network. FTP uses TCP/IP for communication and it works on TCP port 21. FTP works on Client/Server Model where a client requests file from Server and server sends requested resource back to the client.
FTP uses out-of-band controlling i.e. FTP uses TCP port 20 for exchanging controlling information and the actual data is sent over TCP port 21.

The client requests the server for a file. When the server receives a request for a file, it opens a TCP connection for the client and transfers the file. After the transfer is complete, the server closes the connection. For a second file, client requests again and the server reopens a new TCP connection.

Post Office Protocol (POP)

The Post Office Protocol version 3 (POP3) is a simple mail retrieval protocol used by User Agents (client email software) to retrieve mails from mail server.
When a client needs to retrieve mails from server, it opens a connection with the server on TCP port 110. User can then access his mails and download them to the local computer. POP3 works in two modes. The most common mode, the delete mode, is to delete the emails from remote server after they are downloaded to local machines. The second mode, the keep mode, does not delete the email from mail server and gives the user an option to access mails later on mail server.

Hyper Text Transfer Protocol (HTTP)

The Hyper Text Transfer Protocol (HTTP) is the foundation of World Wide Web. Hypertext is well organized documentation system which uses hyperlinks to link the pages in the text documents. HTTP works on client server model. When a user wants to access any HTTP page on the internet, the client machine at user end initiates a TCP connection to server on port 80. When the server accepts the client request, the client is authorized to access web

pages.
To access the web pages, a client normally uses web browsers, who are responsible for initiating, maintaining, and closing TCP connections. HTTP is a stateless protocol, which means the Server maintains no information about earlier requests by clients.
HTTP versions:

HTTP 1.0 uses non persistent HTTP. At most one object can be sent over a single TCP connection.

HTTP 1.1 uses persistent HTTP. In this version, multiple objects can be sent over a single TCP connection.

NETWORK SERVICES

Computer systems and computerized systems help human beings to work efficiently and explore the unthinkable. When these devices are connected together to form a network, the capabilities are enhanced multiple times. Some basic services computer network can offer are:

Directory Services

These services are mapping between name and its value, which can be variable value or fixed. This software system helps to store the information, organize it, and provides various means of accessing it.

Accounting

In an organization, a number of users have their user names and passwords mapped to them. Directory Services provide means of storing this information in cryptic form and make available when requested.

Authentication and Authorization

User credentials are checked to authenticate a user at the time of login and/or periodically. User accounts can be set into hierarchical structure and their access to resources can be controlled using authorization schemes.

DOMAIN NAME SERVICES

DNS is widely used and one of the essential services on which internet works. This system maps IP addresses to domain names, which are easier to remember and recall than IP addresses. Because network operates with the help of IP addresses and humans tend to remember website names, the DNS provides website's IP address which is mapped to its name from the back-end on the request of a website name from the user.

File Services

File services include sharing and transferring files over the network.

File Sharing

One of the reason which gave birth to networking was file sharing. File sharing enables its users to share their data with other users. User can upload the file to a specific server, which is accessible by all intended users. As an alternative, user can make its file shared on its own computer and provides access to intended users.

File Transfer

This is an activity to copy or move file from one computer to another computer or to multiple computers, with help of underlying network. Network enables its user to locate other users in the network and transfers files.

COMMUNICATION SERVICES

Email

Electronic mail is a communication method and something a computer user cannot work without. This is the basis of today's internet features. Email system has one or more email servers. All its users are provided with unique IDs. When a user sends email to other user, it is actually transferred between users with help of email server.

Social Networking

Recent technologies have made technical life social. The computer savvy peoples, can find other known peoples or friends, can connect with them, and can share thoughts, pictures, and videos.

Internet Chat

Internet chat provides instant text transfer services between two hosts. Two or more people can communicate with each other using text based Internet Relay Chat services. These days, voice chat and video chat are very common.

Discussion Boards

Discussion boards provide a mechanism to connect multiple peoples with same interests. It enables the users to put queries, questions, suggestions etc. which can be seen by all other users. Other may respond as well.

Remote Access

This service enables user to access the data residing on the remote computer. This feature is known as Remote desktop. This can be done via some remote device, e.g. mobile phone or home computer.

APPLICATION SERVICES

These are nothing but providing network based services to the users such as web services, database managing, and resource sharing.

Resource Sharing

To use resources efficiently and economically, network provides a mean to share them. This may include Servers, Printers, and Storage Media etc.

Databases

This application service is one of the most important services. It stores data and information, processes it, and enables the users to retrieve it efficiently by using queries. Databases help organizations to make decisions based on statistics.

Web Services

World Wide Web has become the synonym for internet. It is used to connect to the internet, and access files and information services provided by the internet servers.