This chapter introduces general campus switching design considerations and describes modularity in switching designs. It includes the following sections:

- Campus Design Methodology
- Campus Design
- Summary
- Case Study and Simulation Exercise
- Review Questions

CHAPTER 4

Basic Campus Switching Design Considerations

The availability of multigigabit campus switches gives customers the opportunity to build extremely high-performance, high-reliability networks—if they follow correct network design approaches. Unfortunately, some alternative network design approaches can result in a network that has lower performance, reliability, and manageability.

This chapter describes a hierarchical modular design approach called multilayer design. First, it addresses general campus switching design considerations. The differences between Layer 2 (L2) and Layer 3 (L3) switching, and where to use each, are also discussed.

When you finish this chapter, you will be able to understand campus network switch design fundamentals and describe the positioning of switches in campus network modules.

Campus Design Methodology

The multilayer approach to campus network design combines Layer 2 switching with Layer 3 switching to achieve robust, highly-available campus networks. This section discusses the factors you should consider for a Campus local-area network (LAN) design.

Designing an Enterprise Campus

Designing an Enterprise Campus network requires a broad view of the network's overall picture. The network designer must be familiar with both Enterprise Campus design methodologies and Enterprise Campus modules.

Campus design requires an understanding of the organizational network borders (geography) and the existing and planned application traffic flows. Physical characteristics of the network depend on the following criteria:

- Selected transmission media
- The type of technology (switched or shared)
- The type of traffic forwarding (switching) in network devices (Layer 2 or Layer 3)

You should consider the following five factors when deploying the campus network:

- **Network geography**—The distribution of network nodes (for example, host or network devices) and the distances between them significantly affect the campus solution—especially the physical transmission media.
- **Network applications**—In terms of bandwidth and delay, the application requirements place stringent requirements on a campus network solution.
- **Data link layer technology (shared or switched)**—The dedicated bandwidth solution of LAN switching is replacing the traditional approach, in which all devices share the available bandwidth using hubs. The network designer must consider these options, especially when migrating or upgrading existing networks.
- **Layer 2 versus Layer 3 switching**—The network devices and their features determine the network's flexibility, but also contribute to the network's overall delay. Layer 2 switching is based on media access control (MAC) addresses, and Layer 3 switching is based on network layer addresses—usually Internet Protocol (IP) addresses.
- **Transmission media (physical cabling)**—Cabling is one of the biggest long-term investments in network deployment. Therefore, transmission media selection depends not only on the required bandwidth and distances, but also on the emerging technologies that might be deployed over the same infrastructure in the future. The network designer must thoroughly evaluate the cost of the medium (including installation costs) and the available budget in addition to the technical characteristics, such as signal attenuation and electromagnetic interference. Two major cabling options exist: copper-based media (for example, unshielded twisted pair [UTP]) and optical fiber.

The following sections examine these factors.

Network Geography

The location of Enterprise Campus nodes and the distances between them determine the network's geography. When designing the Enterprise Campus network, the network designer's first step is to identify the network's geography. The network designer must determine the following:

- **Location of nodes**—Nodes (end users, workstations, or servers) within an organization can be located in the same room, building, or geographical area.
- **Distances between the nodes**—Based on the location of nodes and the distance between them, the network designer decides which technology should be used, the maximum speeds, and so on. (Media specifications typically include a maximum distance, how often regenerators can be used, and so on.)

The following geographical structures can be identified with respect to the network geography:

- Intra-building
- Inter-building
- Distant remote building
- Distant remote building over 100 km

These geographical structures serve as guides to help determine Enterprise Campus transmission media and the logical modularization of the Enterprise Campus network. The following sections describe these geographical structures.

Intra-Building Structure

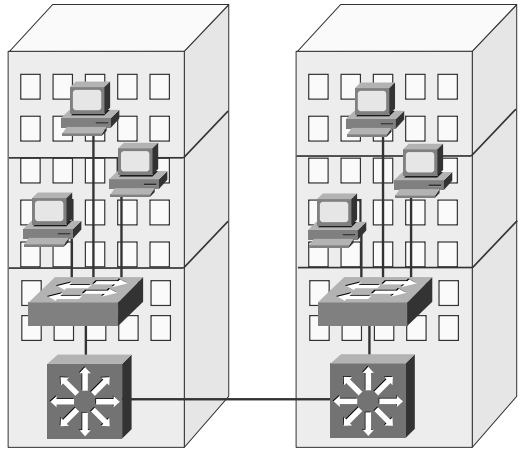
An *intra-building campus network structure* provides connectivity for the end nodes, which are all located in the same building, and gives them access to the network resources. (The access and distribution layers are typically located in the same building.)

User workstations are usually attached to the floor-wiring closet with UTP cables. To allow the most flexibility in the use of technologies, the UTP cables are typically Category 5 (CAT 5) or better. Wiring closets usually connect to the building central switch (distribution switch) over optical fiber. This offers better transmission performances and is less sensitive to environmental disturbances.

Inter-Building Structure

As shown in Figure 4-1, an *inter-building network structure* provides the connectivity between the individual campus buildings' central switches (in the distribution and/or core layers). Typically placed only a few hundred meters to a few kilometers apart, these buildings are usually in close proximity.

Figure 4-1 *Inter-Building Network Structure*



Because the nodes in all campus buildings usually share common devices such as servers, the demand for high-speed connectivity between the buildings is high. To provide high throughput without excessive interference from environmental conditions, optical fiber is the media of choice between the buildings.

Distant Remote Building Structure

When connecting distances that exceed a few kilometers (usually within a metropolitan area), the network designer's most important factor to consider is the physical media. The speed and cost of the network infrastructure depend heavily on the media selection.

Usually, the bandwidth requirements are higher than the physical connectivity options can support. In such cases, the network designer must identify the organization's critical applications and then select the equipment that supports intelligent network services, such as quality of service (QoS) and filtering capabilities that allow optimal use of the bandwidth.

Some companies might own their media, such as fiber or copper lines. However, if the organization does not own physical transmission media to certain remote locations, the Enterprise Network Campus must connect through the Enterprise Edge wide-area network (WAN) module using connectivity options from public service providers (such as metropolitan area network [MAN]).

Network Geography Considerations

Table 4-1 compares the types of connectivity, availability importance, required throughput, and expected cost for each geographical structure.

Table 4-1 *Network Geography Considerations*

Parameter	Intra-building		Inter-building	Distant Remote Building	Distant Over 100 km
Connectivity type	UTP	Fiber	Fiber MM/SM	Fiber SM	Copper/fiber
Availability importance	High	Medium	Medium	Low	Low
Required throughput	Medium	High	High	Medium	Low
Cost	$	$$	$$$	$$$$	$$$$$

MM = Multimode; SM = single-mode

Depending on the distances and environmental conditions that result from the respective geographical scopes, various connectivity options exist—ranging from traditional copper media to fiber-based transmission media.

Typically, availability within a building is very important, and it decreases with distance between buildings. (This is because the physical buildings in the campus often form the core of the campus network; communication to buildings located farther from the core is not as important.)

The throughput requirements increase close to the network's core and close to the sites where the servers reside.

A quick review of Table 4-1 reveals a combination of a high level of availability, medium bandwidth, and a low price for the Enterprise Campus network when all nodes are located in the same building. The cost of transmission media increases with the distance between nodes. A balance between the desired bandwidth and available budget are usually required to keep the cost reasonable; bandwidth is often sacrificed.

Network Application Characterization

Application characterization is the process of determining the characteristics of the network's applications. Network designers should determine which applications are critical to the organization and the network demands of these applications to determine enterprise traffic patterns inside the Enterprise Campus network. This process should result in information about network bandwidth usage and response times for certain applications. These parameters influence the selection of the transmission medium and the desired bandwidth.

Different types of application communication result in varying network demands. The following sections review four types of application communication:

- Client-client
- Client-distributed server
- Client-Server Farm
- Client-Enterprise Edge

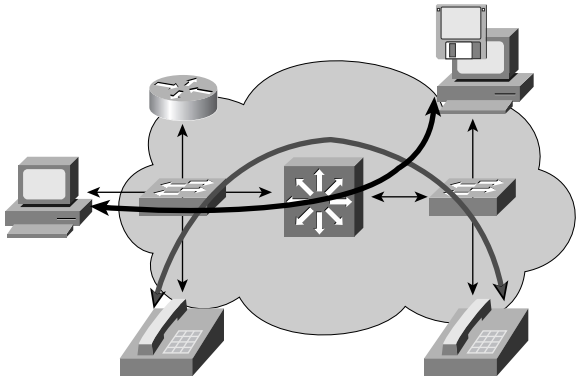
Client-Client Applications

From the network designer's perspective, client-client applications include those applications in which the majority of network traffic passes from one network edge device to another through the organization's network, as shown in Figure 4-2. Typical client-client applications include the following:

- **IP telephony**—Two peers establish communication with the help of a telephone manager workstation; however, the conversation occurs directly between the two peers when the connection is established.
- **File sharing**—Some operating systems (or even applications) require direct access to data on other workstations.

- **Videoconference systems**—This application is similar to IP telephony. However, the network requirements for this type of application are usually higher, particularly bandwidth consumption and QoS requirements.

Figure 4-2 *Client-Client Application*

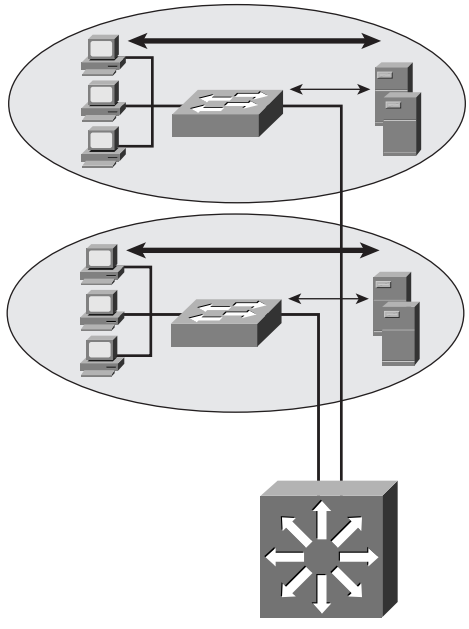


Client-Distributed Server Applications

Historically, clients and servers were attached to a network device on the same LAN segment.

With increased traffic on the corporate network, an organization can decide to split the network into several isolated segments. As shown in Figure 4-3, each of these segments has its own servers, known as *distributed servers*, for its application. In this scenario, servers and users are located in the same virtual LAN (VLAN). Department administrators manage and control the servers. The majority of department traffic occurs in the same segment, but some data exchange (to a different VLAN) can happen over the campus backbone. For traffic passing to another segment, the overall bandwidth requirement might not be crucial. For example, Internet access must go through a common segment that requires less performance than the traffic to the local segment servers.

Figure 4-3 *Client-Distributed Server Application*



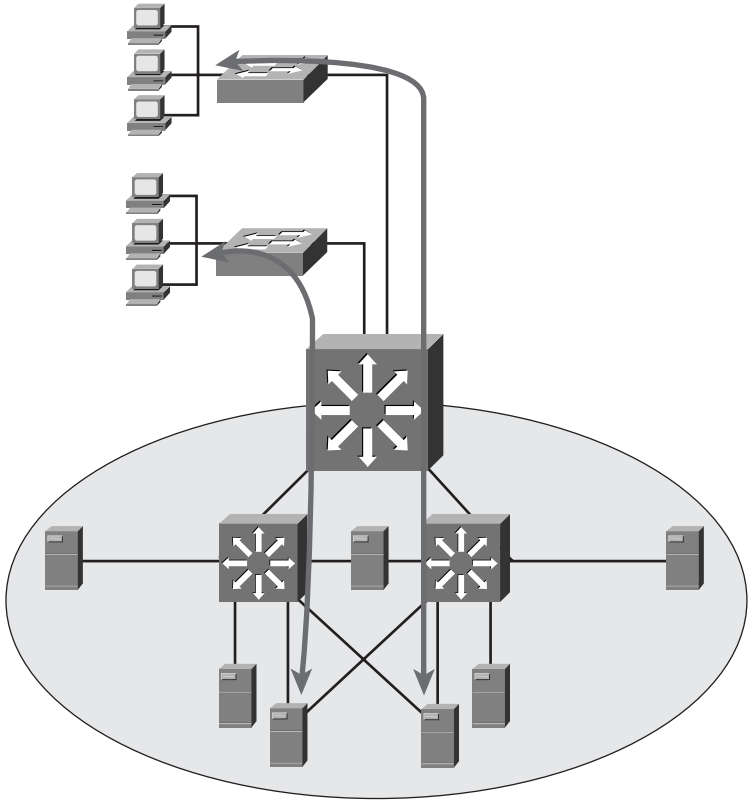
Client-Server Farm Applications

In a large organization, the organizational application traffic passes across more than one wiring closet, or VLAN. Such applications include

- Organizational mail servers (such as Lotus Notes and Microsoft Exchange)
- Common file servers (such as Novell, Microsoft, and Sun)
- Common database servers for organizational applications (such as Sybase, Oracle, and IBM)

A large organization requires its users to have fast, reliable, and controlled access to the critical applications. To fulfill these demands and keep administrative costs down, the solution is to place the servers in a *common Server Farm*, as shown in Figure 4-4. The placement of servers in a Server Farm requires the network designer to select a network infrastructure that is highly resilient (providing security), redundant (providing high availability), and that provides adequate throughput. High-end LAN switches with the fastest LAN technologies, such as Gigabit Ethernet, are typically deployed in such an environment.

Figure 4-4 *Client-Server Farm Application*



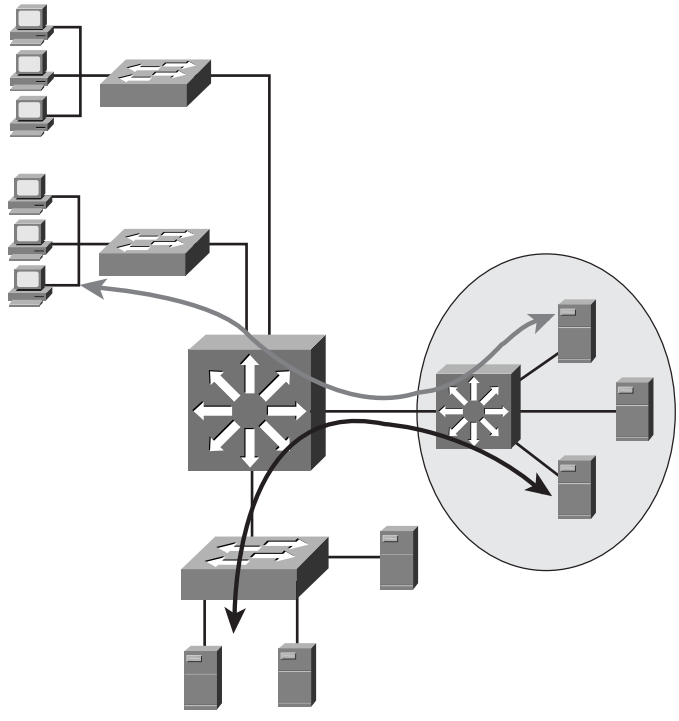
Client-Enterprise Edge Applications

As shown in Figure 4-5, *Client-Enterprise Edge applications* use servers on the Enterprise Edge. These applications exchange data between the organization and its public servers.

The most important communication issue between the Enterprise Campus Network and the Enterprise Edge is not performance, but security. High availability is another important characteristic; data exchange with external entities must be in constant operation. Applications installed on the Enterprise Edge can be crucial to organizational process flow; therefore, any outages can increase costs.

Typical Enterprise Edge applications are based on web technologies. Examples of these application types, such as external mail servers and public web servers, can be found in any organization.

Figure 4-5 *Client-Enterprise Edge Application*



Organizations that support their partnerships through e-commerce applications also place their e-commerce servers into the Enterprise Edge. Communication with these servers is vital because of the two-way replication of data. As a result, high redundancy and resiliency of the network, along with security, are the most important requirements for these applications.

Application Requirements

Table 4-2 compares the types of applications and their requirements for the most important network parameters. The following sections discuss these parameters.

Table 4-2 *Network Application Requirements*

Parameter	Client-Client		Client-Distributed Servers	Client-Server Farm	Client-Enterprise Edge Servers
Connectivity type	Shared	Switched	Switched	Switched	Switched
High availability	Low	Low	Medium	High	High
Total required throughput	Low	Medium	Medium	High	Medium
Total network cost	Low	Low	Medium	High	Medium

Connectivity

The wide use of LAN switching at Layer 2 has revolutionized local-area networking and has resulted in increased performance and more bandwidth for satisfying the requirements of new organizational applications. LAN switches provide this performance benefit by increasing bandwidth and throughput for workgroups and local servers.

NOTE The shared media for client-client (also termed *peer-to-peer*) communication is suitable only in a limited scope, typically when the number of client workstations is very low (for example, in small home offices).

Throughput

The required throughput varies from application to application. An application that exchanges data between users in the workgroup usually does not require a high throughput network infrastructure. However, organizational-level applications usually require a high-capacity link to the servers, which is usually located in the Server Farm.

NOTE Client-client communication, especially in the case of frequent file transfers, could be intensive, and the total throughput requirements can be high.

Applications located on servers in the Enterprise Edge are normally not as bandwidth-consuming (compared to the applications in the Server Farm) but may require high-availability and security features.

High Availability

High availability is a function of the application and the entire network between a client workstation and a server that is located in the network. Although network availability is primarily determined by the network design, the individual components' mean time between failures (MTBF) is a factor. It is recommended that you add redundancy to the distribution layer and the campus.

Cost

Depending on the application and the resulting network infrastructure, the cost varies from low in a client-client environment to high in a highly redundant Server Farm. In addition to the cost of duplicate components for redundancy, costs include the cables, routers, switches, software, and so forth.

Data Link Layer Technologies

Traditionally, network designers had a limited number of hardware options when purchasing a technology for their campus networks. Hubs were used for wiring closets, and routers were used to break the network into logical segments. The increasing power of desktop processors and the requirements of client/server and multimedia applications drove the need for greater bandwidth in traditional shared-media environments. These requirements are prompting network designers to replace hubs with LAN switches.

Key Point: Bandwidth Domains and Broadcast Domains

A *bandwidth domain*, which is known as a *collision domain* for Ethernet LANs, includes all devices that share the same bandwidth. For example, when using switches or bridges, everything associated with one port is a bandwidth domain.

A *broadcast domain* includes all devices that see each other's broadcasts (and multicasts). For example, all devices associated with one router port reside in the same broadcast domain.

Devices in the same bandwidth domain also reside in the same broadcast domain; however, devices in the same broadcast domain can reside in different bandwidth domains.

All workstations residing in one bandwidth domain compete for the same LAN bandwidth resource. All traffic from any host in the bandwidth domain is visible to all the other hosts. In the case of an Ethernet collision domain, two stations can cause a collision by transmitting at the same time. The stations must then stop transmitting and try again at a later time, thereby delaying traffic transmittal.

All broadcasts from any host residing in the same broadcast domain are visible to all other hosts in the same broadcast domain. Desktop protocols such as AppleTalk, Internetwork Packet Exchange (IPX), and IP require broadcasts or multicasts for resource discovery and

advertisement. Hubs, switches, and bridges forward broadcasts and multicasts to all ports. Routers do not forward these broadcasts or multicasts to any ports. In other words, routers block broadcasts (which are destined for all networks) and multicasts; routers forward only unicast packets (which are destined for a specific device) and *directed broadcasts* (which are destined for all devices on a specific network).

Shared Technology

Shared technology using hubs or repeaters is based on all devices sharing a segment's bandwidth. Initially, the entire Ethernet segment was a single common bus—the cable itself. With the introduction of hubs and new structured wiring, the physical network bus topology changed to a star topology. This topology resulted in fewer errors in the network because of the repeaters receiving an electrical signal and boosting the signal before forwarding it to all other segment participants (on all other repeater ports). All devices on all ports of a hub or repeater are on the same bandwidth (collision) domain.

Switched LAN Technology

Switched LAN technology uses the same physical star topology as hubs but eliminates the sharing of bandwidth. Devices on each port of a switch are in different bandwidth (collision) domains; however, all devices are still in the same broadcast domain. The LAN switches provide an efficient way of transferring network frames over the organizational network. In case of a frame error, the switch does not forward the frame as a hub or repeater would.

Comparing Switched and Shared Technologies

Table 4-3 presents some of the most obvious differences and benefits of switched technology compared to shared technology. It uses Fast Ethernet as an example.

Table 4-3 *Switched Versus Shared Fast Ethernet Technologies*

Parameter	Switched	Shared
Bandwidth	>10 Megabits per second (Mbps)	<100 Mbps
Range	From 1 meter	<500 meters
Intelligent services	Yes	No
High availability	Yes	No
Cost	$$	$

Bandwidth

The major drawback of shared technology is that all network devices must compete for the same bandwidth; only one frame flow is supported at a time. Bandwidth in shared technology is limited to the speed on a network segment (in this case, 100 Mbps for Fast Ethernet). Because of collisions, aggregate network bandwidth is less than this.

LAN switching technology supports speeds from Ethernet (10 Mbps) onward and enables multiple ports to simultaneously forward frames over the switch. Thus, the utilized aggregate network bandwidth could be much greater than with shared technology.

NOTE A Layer 3 device separates network segments from each other into different broadcast domains. A traditional network's Layer 3 device was a router; in a modern network, the preference is for a Layer 3 switch.

Range

In a shared network, the network's diameter (the largest distance between two network devices) is constrained by the transmission media's physical characteristics because of the collision detection algorithm—the maximum distance between devices is limited to ensure that no collisions occur. In a shared environment, all devices reside in the same collision domain. The hub improves the frame's physical characteristics but does not check for frame errors. Every station on the segment must compete for resources and be able to detect whether two or more network stations are transmitting at the same time. The Ethernet standard for shared technology defines how long the sending device must possess the bus before it actually sends the data, so collisions can be detected. Because of this time limitation, the length or range of the segment is defined and never reaches more than 500 meters in the best-case scenario.

In a switched environment, devices on each port are in different collision domains. Collision detection is only a concern on each physical segment, and the segments themselves are limited in length. Because the switch stores the entire frame or part of it before forwarding it, the segments do not generate any collisions. The media that is used does not constrain the overall network's diameter.

Intelligent Services

The traditional shared technology is not capable of supporting new network features; this became important with the increasing number of organizational client/server and multimedia applications. LAN switches perform several functions at Layer 3, and even at higher Open System Interconnection (OSI) layers. Modern networks are required to support intelligent network services (such as QoS), security, and management; LAN switches have the ability to support these.

High Availability

Many organizational processes that run on the network infrastructure are critical for the organization's success. Consequently, high availability has become increasingly important. While shared networks do not offer the required capability, the LAN switches do.

Switches can be interconnected with multiple links without creating loops in the network (using the Spanning Tree protocol). Hubs cannot be interconnected with redundant links.

Cost

Considering all the benefits LAN switches offer, you might expect the cost per port to be much higher on switches than on hubs. However, with wide deployment and availability, the price per port for LAN switches is almost the same as it is for hubs or repeaters.

NOTE All of the previously listed factors have mostly eliminated shared technologies; the majority of new networks use only switched technologies. Shared technologies are present in only some parts of existing networks and in smaller home offices.

Layer 2 and Layer 3 Switching Design Considerations

LAN switches have traditionally been only Layer 2 devices. Modern switches provide higher OSI level functionalities and can effectively replace routers in the LAN switched environment. Deploying pure Layer 2 or selecting Layer 3 switches in the enterprise network is not a trivial decision. It requires a full understanding of the network topology and customer demands.

Key Point: Layer 2 Versus Layer 3 Switching

The difference between Layer 2 and Layer 3 switching is the type of information that is used inside the frame to determine the correct output interface. Layer 2 switching forwards frames based on data link layer information (MAC address), while Layer 3 switching forwards frames based on network layer information (such as IP address).

When deciding on the type of LAN switch to use and the features to be deployed into a network, consider the following factors:

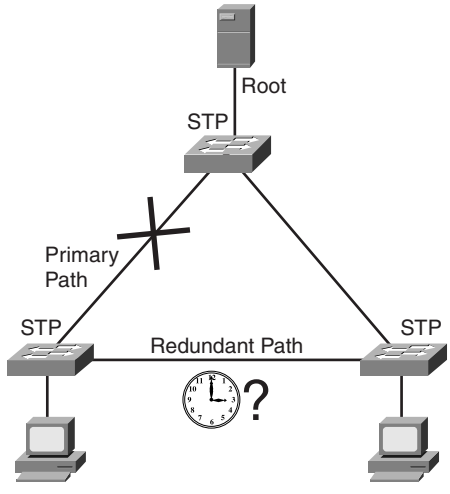
- **Network service capabilities**—The network services the organization requires (QoS, and so on).

- **Size of the network segments**—How the network is segmented, based on traffic characteristics.
- **Convergence times**—The maximum amount of time the network can be unavailable in the event of network outages.

Spanning-Tree Domain Considerations

Layer 2 switches use the Spanning Tree Protocol (STP) to ensure that only one active path exists between two switches. If a physical loop exists (for redundancy), STP puts ports on the switch in *blocking state* (thereby effectively disabling the ports, from a data perspective) to ensure a loop-free network. In the event of a failure, the blocked port is re-enabled (put into a *forwarding state*). An *STP domain* is a set of switches that communicates via STP. STP is illustrated in Figure 4-6.

Figure 4-6 *STP*



STP selects a root switch (or root bridge, according to IEEE 802.1d standard terminology) and determines whether any redundant paths exist. After the switch comes online, it takes up to 50 seconds before the root switch and redundant links are detected. At this time, the switch ports go through the listening and learning states; from there they progress to either the forwarding or blocking state. No ordinary traffic can travel through the network at this time.

NOTE The default STP Forward Delay timer is 15 seconds; it determines how long the port stays in both the listening and learning states (for a total of 30 seconds). The Maximum Age timer defaults to 20 seconds; this is the time during which a switch stores a BPDU before discarding it, and therefore determines when the switch recognizes that a topology change has occurred. The addition of 30 seconds and 20 seconds composes the 50 seconds referred to previously.

When the primary link goes down and the redundant link must be activated, a similar event occurs. The time it takes for a redundant path to be activated depends on whether the failure is direct (a port on the same switch) or indirect (a port on another switch). Direct failures take 30 seconds because the switch bypasses the 20-second Maximum Age timer (and associated Blocking State for the port); from there it moves straight to the listening state (for 15 seconds), and then to the learning state (for 15 seconds). For indirect failures, the switch port must first wait 20 seconds (Maximum Age Timer) before it can transition to the listening state and then the learning state, for a total of 50 seconds. Thus, when a link fails, up to 50 seconds might pass before another link becomes available.

Cisco has implemented several features that have improved STP convergence. Recent standardization efforts have also proposed some new enhancements to the STP. Following is a brief description of the STP enhancements that result in faster convergence; this convergence is comparable to Layer 3 convergence and, in some instances, even exceeds it.

- **PortFast**—Used for ports in which end-user stations and/or servers are directly connected. When PortFast is enabled, there is no delay in passing traffic because the switch immediately puts the port in the forwarding state (skipping the listening and learning states). Two additional measures that prevent potential STP loops are associated with the PortFast feature:
  - **Bridge Protocol Data Unit (BPDU) Guard**—PortFast transitions the port into STP forwarding mode immediately upon linkup. Since the port still participates in STP, the potential of STP loop exists (if some device attached to that port also runs STP). The BPDU guard feature enforces the STP domain borders and keeps the active topology predictable. If the port receives a BPDU, the port is transitioned into *errdisable state* (meaning that it was disabled due to an error) and an error message is reported.

NOTE Additional information regarding the errdisable state is available in *Recovering From errDisable Port State on the CatOS Platforms*, at www.cisco.com/en/US/tech/tk389/tk214/technologies_tech_note09186a0080093dcb.shtml.

  - **BPDU Filtering**—This feature allows the user to block PortFast-enabled nontrunk ports from transmitting BPDUs. Spanning tree does not run on these ports.

- **UplinkFast**—If the link to the root switch goes down and the link is directly connected to the switch, UplinkFast enables the switch to put a redundant path (port) into active state within a second.
- **BackboneFast**—If a link on the way to the root switch fails but is not directly connected to the switch, BackboneFast reduces the convergence time from 50 seconds to between 20 and 30 seconds. When this feature is used, it must be enabled on all switches in the STP domain.

In addition to features that enable faster convergence of the STP, features exist that prevent errors from resulting in unpredictable STP topology changes that could lead to STP loops. These features include the following:

- **STP Loop Guard**—When one of the blocking ports in a physically redundant topology stops receiving BPDUs, usually STP creates a potential loop by moving the port to forwarding state. With the STP Loop Guard feature enabled and if a blocking port no longer receives BPDUs, that port is moved into the STP loop-inconsistent blocking state instead of the listening/learning/forwarding state. This feature avoids loops in the network that result from unidirectional or other software failures.
- **BPDU Skew Detection**—This feature allows the switch to keep track of late-arriving BPDUs (by default, BPDUs are sent every 2 seconds) and notify the administrator via syslog messages. Skew detection generates a report for every port on which BPDU has ever arrived late (this is known as *skewed*). Report messages are rate-limited (one message every 60 seconds) to protect the CPU.
- **Unidirectional Link Detection (UDLD)**—If the STP process that runs on the switch with a blocking port stops receiving BPDUs from its upstream (designated) switch on that port, STP creates a forwarding loop or STP loop by eventually aging out the STP information for this port and moving it to the forwarding state. The UDLD is a Layer 2 protocol that works with the Layer 1 mechanisms to determine a link's physical status. If the port does not see its own device/port ID in the incoming UDLD packets for a specific duration of time, the link is considered unidirectional from the Layer 2 perspective. Once UDLD detects the unidirectional link, the respective port is disabled and the error message is generated.

Although spanning tree was previously considered to have very slow convergence (up to 50 seconds), the latest standard enhancements render its convergence comparable to (or even exceeding) that of routing protocols. The following enhancements are useful in environments that contain several VLANs:

- **Rapid STP (RSTP, defined in IEEE 802.1W)**—RSTP provides rapid convergence of the spanning tree by assigning port roles and determining the active topology. The RSTP builds upon the IEEE 802.1d STP to select the switch with the highest switch priority as the root switch and then assigns the port roles (root, designated, alternate, backup, and disabled) to individual ports. These roles assist in rapid STP convergence, which can be extremely fast (within a second) because of the topology knowledge.

- **Multiple STP (MSTP, sometimes referred to as MISTP [Multiple Instances of STP], defined in IEEE 802.1S)**—MSTP uses RSTP for rapid convergence by enabling several (topologically identical) VLANs to be grouped into a single spanning tree instance, with each instance including a spanning tree topology that is independent of other spanning tree instances. This architecture provides multiple forwarding paths for data traffic, enables load balancing, and reduces the number of spanning tree instances that are required to support a large number of VLANs.

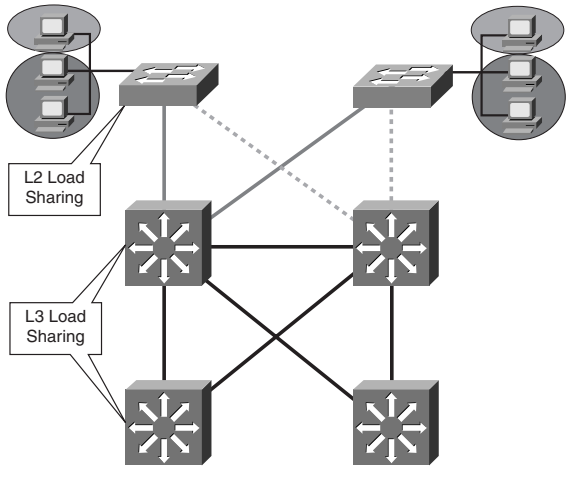
Load Sharing Guidelines

Layer 2 and Layer 3 switches handle load sharing differently, as described in the following sections.

Layer 2 Load Sharing

Because Layer 2 switches are aware of only MAC addresses, they cannot perform any intelligent load sharing. In an environment characterized by multiple VLANs per access switch and more than one connection to the uplink switch, the solution is to put all uplink connections into trunks (Inter-switch link [ISL] or 802.1q). Each trunk carries all VLANs; however, without additional configuration, the STP protocol disables all nonprimary uplink ports. This configuration can result in a bandwidth shortage because the traffic for all the VLANs passes through the same link. To overcome this problem, the STP parameters must be configured to carry some VLANs across one uplink and the rest of the VLANs across the other uplink. For example, one uplink could be configured to carry the VLANs with odd numbers, while the other uplink is configured to carry the VLANs with even numbers. The top of Figure 4-7 illustrates this situation.

Figure 4-7 *Layer 2 Versus Layer 3 Load Sharing*



Layer 3 Load Sharing

Layer 3-capable switches can perform load sharing based on IP addresses. As illustrated in the lower portion of Figure 4-7, most modern Layer 3 devices with load sharing capability can balance the load per packet or per destination-source IP pair.

The advantage of Layer 3 IP load sharing is that links are used more proportionately than with Layer 2 load sharing, which is based on VLANs only. For example, the traffic in one VLAN can be very heavy while the traffic in another VLAN is very low; in this case, per-VLAN load sharing by using even and odd VLANs is not appropriate. Due to the dynamic nature of organizational applications, Layer 3 load sharing is more appropriate. Layer 3 allows for dynamic adaptation to link utilization and depends on the routing protocol design. Layer 3 switches also support Layer 2 load sharing, so they can still apply per-VLAN load sharing while connected to other Layer 2 switches.

Layer 2 Versus Layer 3 Switching

Table 4-4 compares Layer 2 and Layer 3 switching with respect to various campus network features. Considerations for deployment include

- Pure Layer 2 switching throughout the network
- Various combinations of Layer 2 and Layer 3 switching, including
  - Layer 3 switching in the distribution layer only
  - Layer 3 switching in the distribution and core layers
- Layer 3 switching throughout the network

Table 4-4 *Layer 2 Versus Layer 3 Switching*

Parameter	Layer 2 Everywhere	Layer 3 in Distribution Only	Layer 3 in Core and Distribution	Layer 3 Everywhere
Policy domain	Layer 2 Access Control List (ACL) and QoS	Layer 2 and Layer 3 ACL and QoS	Layer 2 and Layer 3 ACL and QoS	Layer 2 and Layer 3 ACL and QoS
Load sharing	Per VLAN	Per VLAN Per destination	Per VLAN Per destination	Per VLAN Per destination
Failure domain	VLAN	Access, core	Access	Segment
Convergence	STP	Distribution: Routing protocol hold-timer (quick) Other: STP	Core and distribution: Routing protocol hold-timer (quick) Access: STP	Routing protocol hold-timer (quick)
Cost	$→	$$→	$$$→	$$$$

The following sections elaborate on the features in Table 4-4.

Policy Domain

The *policy domain* is the scope of the network that is affected by a certain policy. A *network policy* is a formal set of statements that define how network resources are allocated among devices. In addition to selected hosts or applications, the policies can be applied to individual users, groups, or entire departments. For example, policies can be based on the time of day or client authorization priorities. Network managers implement policies and policy statements and store them in a policy repository or on the device itself. The devices then apply the configured policies to network resources.

The size of the policy domain depends on the switching layer and on the mechanisms for policy implementation. In pure Layer 2 switching, the policy domain overlaps with the switching domain's boundaries; Layer 3 switching offers much more flexibility. In Layer 2 switching, the access control lists (ACLs) and various QoS mechanisms can only be applied to switched ports and MAC addresses; in the Layer 3 switching, the ACL and QoS mechanisms are extended to IP addresses, or even applications (for example, using Transmission Control Protocol [TCP] and User Datagram Protocol [UDP] ports).

Load Sharing

When multiple links exist, they can be used for redundancy and/or traffic load sharing. As discussed in the "Load Sharing Guidelines" section of this chapter, Layer 2 switches only offer load sharing by distributing VLANs across different uplink ports. Layer 3 switches, however, can perform load sharing between ports based on IP destinations.

Failure Domain

A *failure domain* defines the scope of the network that is affected by network failures. In a Layer 2-switched domain, a misconfigured or malfunctioning workstation can introduce errors that impact or disable the entire domain. Problems of this nature are often difficult to localize.

A failure domain is

- Bounded by Layer 3 switching
- Bounded by the VLAN when Layer 2 switching is deployed in an entire campus

Convergence

As discussed in the "Spanning-Tree Domain Considerations" section of this chapter, loop prevention mechanisms in a Layer 2 topology cause the STP to take between 30 and 50

seconds to converge. To eliminate STP convergence issues in the campus backbone, all the links connecting backbone switches must be routed links, not VLAN trunks. This also limits the broadcast and failure domains.

In the case where the Layer 3 switching is deployed everywhere, convergence is within seconds (depending on the routing protocol implemented) because all the devices detect their connected link failure immediately and act upon it promptly (sending respective routing updates).

In a mixed Layer 2 and Layer 3 environment, the convergence time not only depends on the Layer 3 factors (including routing protocol timers such as hold-time and neighbor loss detection), but also on the STP convergence.

Using Layer 3 switching in a structured design reduces the scope of spanning tree domains. It is common to use a routing protocol, such as Enhanced Interior Gateway Protocol (EIGRP) or Open Shortest Path First (OSPF), to handle load balancing, redundancy, and recovery in the backbone.

Cost

The cost of deploying Layer 3 switching in comparison to Layer 2 switching increases with the scope of Layer 3 switching deployment. Layer 3 switches are more expensive than their Layer 2 counterparts; for example, Layer 3 functionality can be obtained by adding cards and software to a modular Layer 2 switch.

Transmission Media

An Enterprise Campus can use various physical media to interconnect devices.

Selecting the type of cable is an important consideration when deploying a new network or upgrading an existing one. Cabling infrastructure represents a long-term investment—it is usually installed to last for ten years or more. In addition, even the best network equipment does not operate as expected with poorly chosen cabling.

A network designer must be aware of physical media characteristics because they influence the maximum distance between devices and the network's maximum transmission speed.

Twisted-pair cables (copper) and optical cables (fiber) are the most common physical transmission media used in modern networks.

Unshielded Twisted-Pair (UTP) Cables

UTP consists of four pairs of isolated wires that are wrapped together in plastic cable. No additional foil or wire is wrapped around the core wires (thus, they are *unshielded*). This makes these wires less expensive, but also less immune to external electromagnetic

influences than shielded cables. UTP is widely used to interconnect workstations, servers, or other devices from their network interface card (NIC) to the network connector at a wall outlet.

The characteristics of twisted-pair cable depend on the quality of their material. As a result, twisted-pair cables are sorted into categories. Category 5 or greater is recommended for speeds of 100 megabits per second (Mbps) or higher. Because of the possibility of signal attenuation in the wires and carrier detection, the maximum cable length is usually limited to 100 meters . For example, if one PC starts to transmit and another PC is more than 100 meters away, the second PC might not detect the signal on the wire and therefore start to transmit, causing a collision on the wire.

One of the frequent considerations in the cabling design is electromagnetic interference. Due to high susceptibility to interference, UTP is not suitable for use in environments with electromagnetic influences. Similarly, UTP is not appropriate for environments that can be affected by the UTP's own interference.

NOTE Some security issues are also associated with electromagnetic interference—it is easy to eavesdrop on the traffic carried across UTP because these cables emit electromagnetic interference.

Optical Cables

Typical requirements that lead to the selection of optical cable as a transmission media include distances longer than 100 meters, and immunity to electromagnetic interference. There are different types of optical cable; the two main types are multimode (MM) and single-mode (SM).

Both MM and SM optical cable have lower signal losses than a twisted pair cable; therefore, optical cables automatically enable longer distances between devices. However, fiber cable has precise production and installation requirements, resulting in a higher cost than twisted pair cable.

Multimode fiber is optical fiber that carries multiple light waves or modes concurrently, each at a slightly different reflection angle within the optical fiber core. Because modes tend to disperse over longer lengths (modal dispersion), MM fiber transmission is used for relatively short distances. Typically, light emitting diodes (LEDs) are used with MM fiber. The typical diameter of an MM fiber is 50 or 62.5 micrometers.

Single-mode (also known as *monomode*) *fiber* is optical fiber that carries a single wave (or laser) of light. Lasers are typically used with SM fiber. The typical diameter of an SM fiber core is between 2 and 10 micrometers.

Copper Versus Fiber

Table 4-5 presents some of the critical parameters that influence the network transmission medium selection.

Table 4-5 *Copper Versus Fiber Media*

Parameter	Copper	Fiber
Bandwidth	Ethernet: <1 gigabits per second (Gbps) LRE: <15 Mbps	<10 Gbps
Range	Ethernet: <100 m LRE: <1.5km	MM: 550 m* SM: <100 km*
Deployment area	Wiring closet	Inter-node and inter-building
Other considerations	Interference, grounding	Coupling loss
Installation cost	$	$$$

* When using Gigabit Ethernet

NOTE Table 4-5 lists Ethernet as a technology; this includes Ethernet, Fast Ethernet, and Gigabit Ethernet. Long Reach Ethernet (LRE) is also listed. This latter technology is Cisco proprietary and runs on voice-grade copper wires; it allows higher distances than traditional Ethernet and is used as an access technology in WANs. Chapter 5, "Designing WANs," further describes LRE.

The following sections elaborate on the parameters in Table 4-5.

Bandwidth

The *bandwidth parameter* indicates the required bandwidth in a particular segment of the network, or the connection speed between the nodes inside or outside the building.

Range

The *range parameter* is the maximum distance between network devices (such as workstations, servers, printers, and IP phones) and network nodes, and between network nodes.

Table 4-6 summarizes the bandwidth and range characteristics of the transmission media types.

Table 4-6 *Transmission Media Types Bandwidth and Range Characteristics*

Parameter	Twisted-Pair	MM Fiber	SM Fiber
Distance (range)	Up to 100 meters	Up to 2 kilometers (km) (Fast Ethernet) Up to 550 m (Gigabit Ethernet)	Up to 40 km Up to 100 km (Gigabit Ethernet)
Speed (bandwidth)	Up to 1 Gpbs	Up to 1 Gbps	10 Gbps
Cost	Cheap to install	Moderate	High price

Copper cables are typically used for connectivity of network devices to the wiring closet where

- Distances are less than 100 meters
- Speeds of 100 Mbps are satisfactory
- Cost must be kept within reasonable limits

NOTE Fast EtherChannel (FEC) and Gigabit EtherChannel solutions group several parallel links between LAN switches into a channel that is seen as a single link from the Layer 2 perspective. Two protocols have been introduced for automatic EtherChannel formation: the Port Aggregation Control Protocol (PagP), which is Cisco proprietary, and the Link Aggregation Control Protocol (LACP), which is standardized and defined in IEEE 802.3ad.

Deployment Area

Deployment area indicates whether wiring is required for wiring closet only (where users access the network), for inter-node, or even for inter-building connections.

Connection from the wiring closet to the building central node can use UTP. As for most inter-node and especially inter-building connections, MM, or even SM, fiber is probably needed if there are high-speed requirements.

Other Considerations

When deploying UTP in an area with high electrical or magnetic interference—for example, in an industrial environment—you must pay special attention to media selection. In such environments, the disturbances might interfere with data transfer and therefore

result in an increased number of frame errors. Electrical grounding can isolate some external disturbance, but the wiring increases the costs. Fiber optic installation is the only reasonable solution for such networks.

Optical fiber requires a precise technique for cable coupling. Even a small deviation from the ideal position of optical connectors can result in either a loss of signal or a large number of frame losses. Careful attention during optical fiber installation is imperative because of the traffic's high sensitivity to coupling misalignment. In environments where the cable does not consist of a single fiber from point to point, coupling is required and loss of signal can easily occur.

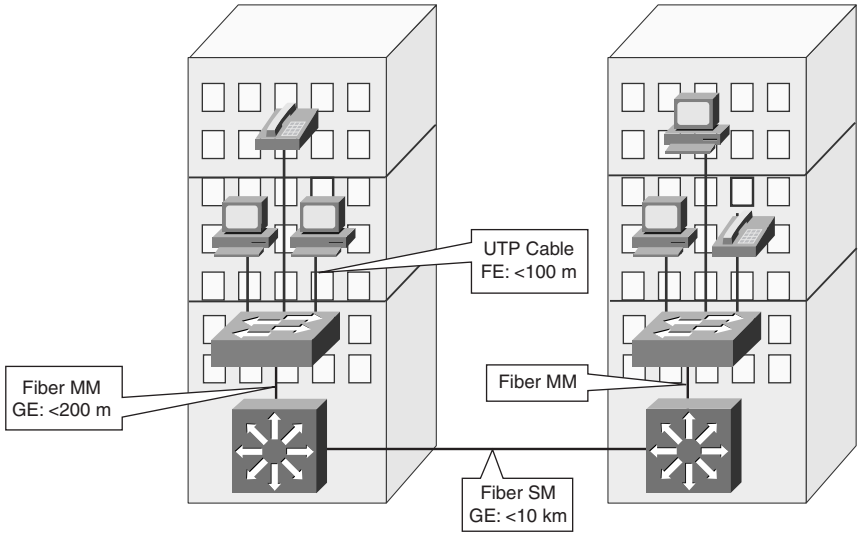
Installation Cost

Along with the cost of the medium, you must also seriously consider installation cost. Installation costs are significantly higher than UTP installation costs because of strict requirements for optical cable coupling.

Cabling Example

Figure 4-8 illustrates a typical campus network structure. End devices such as workstations, IP phones, and printers are no more than 100 m away from the LAN switch. UTP wiring can easily handle the required distance and speed; it is also easy to set up, and the price/performance ratio is reasonable.

Figure 4-8 *A Campus Network Uses Many Different Types of Cables*



Optical fiber cables handle higher speeds and distances that can be required among switch devices. MM optical cable is usually satisfactory inside the building. Depending on distance, organizations use MM or SM optical for inter-building communication cable. If the distances are short (up to 500 m), MM fiber is a more reasonable solution for speeds up to 1 Gbps.

However, an organization can install SM fiber if its requirements are for longer distances, or if they are planning for future higher speeds (for example, 10 Gbps). The current specification provides Gigabit Ethernet connectivity on SM fiber up to 5 km; however, Cisco has already provided modules that support connectivity up to 10 km, and even up to 100 km.

NOTE Selecting the less expensive type of fiber might satisfy a customer's current need, but this fiber might not meet the needs of future upgrades or equipment replacement. Replacing cable can be very expensive. Planning with future requirements in mind might result in higher initial costs, but ultimately lower costs.

Campus Design

Campus building blocks are comprised of multilayer devices that connect to the campus backbone. A building design is appropriate for a building-sized network that contains several thousand networked devices; a campus design is appropriate for a large campus that consists of many buildings. To scale from a building model to a campus model, network designers must add a campus backbone between buildings.

This section discusses advanced network traffic considerations and building design using Layer 2 and Layer 3 switching in the access and distribution layers. It describes traffic patterns, multicast traffic, and QoS, and uses both Layer 2 and Layer 3 technologies to discuss campus backbone design. Finally, we investigate server placement within the campus and present guidelines for connectivity to the rest of the enterprise network.

Introduction to Enterprise Campus Design

As discussed in Chapter 3, "Structuring and Modularizing the Network," the Enterprise Campus network can be divided into the following modules:

- **Campus Infrastructure**—This module contains the following submodules:
  - **Building Access**—Aggregates end user connections and provides access to the network.
  - **Building Distribution**—Provides aggregation of access devices and connects them to the campus backbone.
  - **Campus Backbone**—Interconnects the building distribution submodules with the Edge Distribution module and provides high-speed transport.

- **Server Farm**—Connects the servers to the enterprise network and manages the campus server availability and traffic load balancing.
- **Edge Distribution**—Connects the Enterprise Edge applications to the network campus. Security is the main consideration in this module.
- **Network Management**—The Network Management module requirements are similar to those for the Server Farm module, with the exception of the bandwidth requirement. The Network Management module typically does not require high bandwidth.

This section identifies major requirements for designing campus networks within these modules.

Enterprise Campus Module Requirements

As shown in Table 4-7, each Enterprise Campus module has different requirements. For example, this table illustrates how modules that are located closer to the users require a higher degree of scalability. This means that the network designer must consider an option for expanding the Campus network easily in the future, without redesigning the complete network. For example, adding new workstations to a network should result in neither high investment cost nor performance degradations.

Table 4-7 *Enterprise Campus Design Requirements*

Requirement	Building Access		Building Distribution	Campus Backbone	Server Farm	Edge Distribution
Technology	Shared	Layer 2 switched	Layer 2 and 3 switched	Layer 2 and 3 switched	Layer 3 switched	Layer 3 switched
Scalability	High	High	Medium	Low	Medium	Low
High availability	Low	Medium	Medium	High	High	Medium
Performance	Low	Low	Medium	High	High	Medium
Cost per port	Low	Low	Medium	High	High	Medium

The end user usually does not require high performance and high availability, but they are crucial to the campus backbone—especially the Server Farm module.

The price per port increases with increased performance and availability. The campus backbone and Server Farm require a guarantee of higher throughput so they can handle all traffic flows and not introduce additional delays or drops to the network traffic.

The Edge Distribution module does not require the same performance as in the campus backbone. However, it can require other features and functionalities that increase the overall cost.

Enterprise Campus Design Considerations

Designing an Enterprise Campus means not only dividing the network into modules, but also optimizing performance and the cost of each module while providing scalability and high availability. Before designing a campus network, you must take the following considerations relating to network traffic into account:

- **Application traffic patterns**—Identify the organizational traffic flows. This includes the type of traffic and its bandwidth requirements and traffic patterns.
- **Multicast traffic**—Identify the features that constrain multicast streams to the relevant ports. If present in the Enterprise Campus network and incorrectly designed, multicast traffic can use a great amount of bandwidth.
- **Delay sensitive traffic**—Identify and incorporate the appropriate QoS mechanisms to manage the diverse requirements for delay and delay variations.

As Table 4-8 shows, the Enterprise Campus can be built on either a shared or switched (Layer 2 or Layer 3) foundation technology. In the building access layer, workstations with low demand can be connected via shared technology; however, this option is only suitable for some small (home) offices that have a few devices without any special bandwidth requirements. Where higher speeds are required, shared technology is not appropriate, and LAN switching is the only option. The remaining consideration is whether to use Layer 2 or Layer 3 switching technology.

Table 4-8 *Enterprise Campus Design Decisions*

Requirement	Building Access		Building Distribution	Campus Backbone	Server Farm	Edge Distribution
Technology	Shared	Layer 2 switched	Layer 2 and 3 Switched	Layer 2 and 3 switched	Layer 3 switched	Layer 3 switched
Application traffic	Distant	Local/ distant	Distant	Distant	Local/ distant	Distant
Multicast traffic aware	No	Layer 2 limited	Yes	Yes	Yes	Yes
QoS (delay sensitive) traffic support	No	Queuing/marking per port Marking per application				

Consideration of the applications and traffic is required to ensure that the appropriate equipment for the individual modules is selected. Application traffic patterns, multicast traffic, and QoS are important network design issues.

Layer 2 switches usually support multicast and QoS features, but with limited capability. A Layer 3 switch, or in the case of IP multicast, at least a so-called *Layer 3-aware switch*, might be required.

A Layer 2 multicast-aware switch that works closely with the Layer 3 device (router) can distinguish which hosts belong to the multicast stream and which do not. Thus, the Layer 2 switch can forward the multicast stream to only selected hosts.

Layer 2 QoS support is usually limited to port marking capability and queuing on only uplink trunk ports, especially on low-end switches. Layer 2 switches are usually incapable of marking or queuing based on the Layer 3 parameters of packets. However, several recent platforms have added support for Layer 2, Layer 3, and Layer 4 class of service (CoS) and type of service (ToS) packet marking and policing.

The following sections examine network traffic patterns, multicast traffic, and QoS considerations in the Enterprise Campus modules.

Network Traffic Patterns

Campus traffic patterns are generally categorized as local (within a segment or submodule) or distant (passing several segments and crossing the module boundaries).

Network traffic patterns have changed through the years. The characteristic of traditional campus networks was 80 percent local traffic and 20 percent distant traffic; this is known as the *80/20 rule*. In modern campus networks, the ratio is closer to 20/80 because the servers are no longer present in the workgroup, but are instead placed separately in the Server Farm. The 20/80 ratio results in a much higher load on the backbone because the majority of the traffic from client workstations to the servers passes through the backbone.

80/20 Rule in the Campus

When designing a switched campus, network designers ensure that each switched segment corresponds to a workgroup. By placing the workgroup server in the same segment as its clients, most of the traffic can be contained. The 80/20 rule refers to the goal of containing at least 80 percent of the traffic within the local segment.

The campus-wide VLAN model is highly dependent upon the 80/20 rule. If 80 percent of the traffic is within a workgroup (VLAN), 80 percent of the packets flowing from the client to the server are switched locally.

The conventional 80/20 rule underlies traditional network design models. With the campus-wide VLAN model, the logical workgroup is dispersed across the campus, but is still organized so that 80 percent of traffic is contained within the VLAN. The remaining 20 percent of traffic leaves the network through a router.

20/80 Rule in the Campus

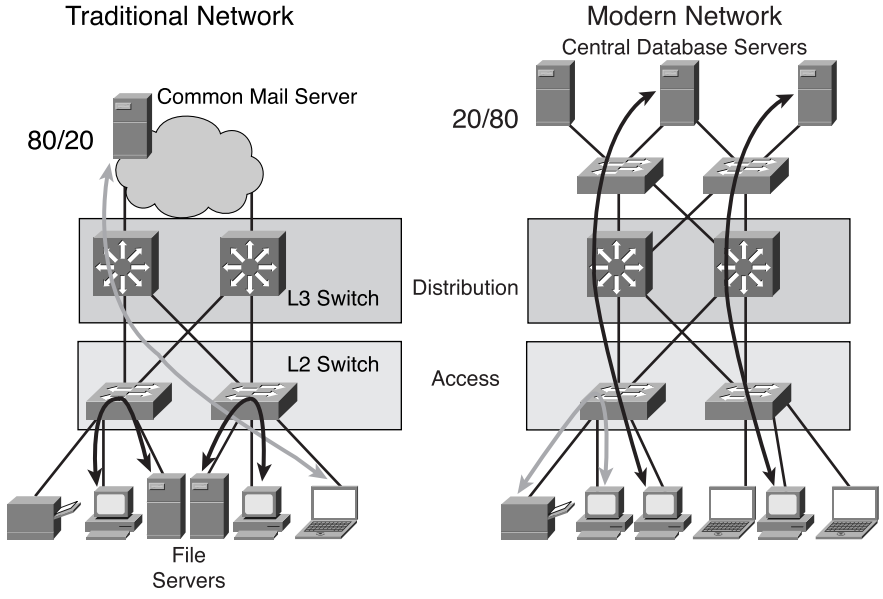
Many new and existing applications currently use distributed data storage and retrieval. The traffic pattern is moving toward what is now referred to as the *20/80 rule*. With the 20/80 rule, only 20 percent of traffic is local to the workgroup LAN, and 80 percent of the traffic leaves the workgroup.

In a traditional network design, only a small amount of traffic passes through the Layer 3 devices. Because performance was not an issue, these devices have traditionally been routers. Modern enterprise networks utilize servers that are located in Server Farms or in the enterprise edge. With an increasing amount of traffic from clients to distant servers, performance requirements are higher in the building distribution and campus backbone. Therefore, devices that have a very high speed of Layer 3 processing are necessary; these devices are Layer 3 switches.

Network Traffic Pattern Example

Figure 4-9 illustrates examples of the 80/20 and 20/80 rules in a campus network.

Figure 4-9 *Traffic Patterns in Traditional and Modern Networks*



Company A, shown on the left side of Figure 4-9, has several independent departments. Each department has its own VLAN, in which the servers and printers are located. File transfers from other department servers or workstations are necessary only occasionally. This traffic must pass the distribution layer, which is represented by the Layer 3 switch. The only common resource the departments use is the mail server, which is located in the corporate network's core.

Company B, shown on the right side of Figure 4-9, also has several departments; however, they use common resources. Not only do they use file servers from their own department, but they also use services from common data storage, such as an Oracle database. This type of configuration requires a higher-performance Layer 3 switch on the distribution layer. The access layer switch (Layer 2) concentrates users into their VLANs. The servers on the other side of the network are also organized into groups and are connected to Layer 2 switches. Distribution layer switches in the middle enable fast, reliable, and redundant communication among the groups on both sides of the network. Figure 4-9 illustrates that the majority of the communication takes place between servers and users, and only a small amount of traffic is switched inside the group.

Multicast Traffic Considerations

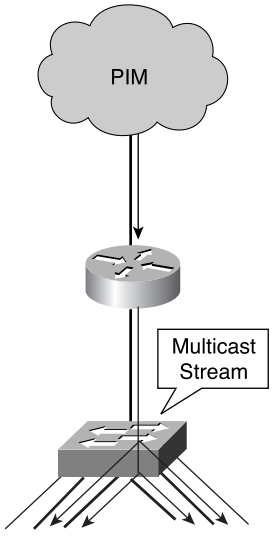
IP multicast is a bandwidth-conserving technology that reduces traffic by simultaneously delivering a single stream of information to potentially thousands of corporate recipients.

Videoconferencing, corporate communications, distance learning, distribution of software, stock quotes, and news are some applications that take advantage of the multicast traffic stream. IP multicast delivers source traffic to multiple receivers.

IP multicast is based on the concept of a multicast group. Any group of receivers can express an interest in receiving a particular data stream. This group does not require any physical or departmental boundaries; rather, the hosts can be located anywhere on the corporate network. Hosts that are interested in receiving data that flows to a particular group must join the group using the Internet Group Management Protocol (IGMP).

Figure 4-10 illustrates a typical situation with IP multicast. Multicast-enabled routers ensure that traffic is delivered properly by using one of the multicast routing protocols, such as Protocol Independent Multicast (PIM). The router forwards the incoming multicast stream to the switch port.

Figure 4-10 *Multicast Traffic Handled by Router*



However, the default behavior for a Layer 2 switch is to forward all multicast traffic to every port that belongs to the same VLAN on the switch (a behavior known as *flooding*). This behavior defeats the purpose of the switch, which is to limit the traffic to only the ports that must receive the data.

NOTE Support for broadcast and multicast suppression is available on several switched platforms. The suppression is done with respect to the incoming traffic rate and is either bandwidth-based or measured in packets per second. The threshold can be set to any value between 0 and 100 percent (or as a number of packets when packet-based suppression is turned on). When the data on the port exceeds the threshold, the switch suppresses further activity on the port for the remainder of the 1-second period.

Static entries can sometimes be set to specify which ports should receive the multicast traffic. Dynamic configuration of these entries simplifies the switch administration.

Several methods exist for Cisco switches to deal efficiently with multicast in a Layer 2 switching environment. Following are the most common methods:

- **Cisco Group Management Protocol (CGMP)**—CGMP allows switches to communicate with a router to determine whether any of the users attached to them are part of a multicast group. The multicast receiver registration is accepted by the router

(using the IGMP) and communicated via CGMP to the switch; the switch adjusts its forwarding table accordingly. CGMP is a Cisco proprietary solution that is implemented on all Cisco LAN switches.

- **IGMP snooping**—With IGMP snooping, the switch intercepts multicast receiver registrations and adjusts its forwarding table accordingly. IGMP snooping requires the switch to be Layer 3-aware because IGMP is a network layer protocol. Typically, the IGMP packet recognition is hardware-assisted.

NOTE Additional methods for addressing the problem of multicast frames in a switched environment include the Generic Multicast Registration Protocol (GMRP) and the Router-Port Group Management Protocol (RGMP). GMRP, which is used between the switch and the host, is not yet widely available. RGMP is a Cisco solution for router-only multicast interconnects in a switched environment. (More information on RGMP is available in *Configuring RGMP,* at www.cisco.com/en/US/products/hw/switches/ps708/products_configuration_guide_chapter09186a008007e6f8.html.)

QoS Considerations for Delay-sensitive Traffic

A Campus Network transports many types of applications and data, including high-quality video and delay-sensitive data (such as real-time voice). Bandwidth-intensive applications stretch network capabilities and resources, but they can also enhance many business processes. Networks must provide secure, predictable, measurable, and sometimes guaranteed services. Achieving the required QoS by managing delay, delay variation (jitter), bandwidth, and packet loss parameters on a network can be the key to a successful end-to-end business solution. QoS mechanisms are techniques that are used to manage network resources.

The assumption that a high-capacity, nonblocking switch with multigigabit backplanes never needs QoS is incorrect. Most networks or individual network elements are oversubscribed. In fact, it is easy to create scenarios in which congestion potentially occurs and that therefore require some form of QoS. Uplinks from the access layer to the distribution layer, or from the distribution layer to the core, most often require QoS. The sum of the bandwidths on all ports on a switch where end devices are connected is usually greater than that of the uplink port. When the access ports are fully used, congestion on the uplink port is unavoidable.

Depending on traffic flow and uplink oversubscription, bandwidth is managed with QoS mechanisms on the access, distribution, or even core switches.

QoS Categories

Layer 2 QoS is similar to Layer 3 QoS, which Cisco IOS software implements. You can configure the following four QoS categories on LAN switches:

- **Classification and marking**—Packet classification features allow the partitioning of traffic into multiple priority levels, or classes of service. These features inspect the information in the frame header (Layer 2, Layer 3, and Layer 4) and determine the frame's priority. *Marking* is the process of changing a frame's CoS setting (or priority).
- **Scheduling**—Scheduling is the process that determines the order in which queues are serviced. CoS is used on Layer 2 switches to assist in the queuing process. Layer 3 switches can also provide QoS scheduling; Layer 3 IP QoS queue selection uses the IP DiffServ Code Point (DSCP) or the IP packet's IP precedence field.

NOTE For more information on DSCP and IP precedence, refer to the Cisco *Implementing Quality of Service Policies with DSCP* document at www.cisco.com/en/US/tech/tk543/tk757/technologies_tech_note09186a00800949f2.shtml.

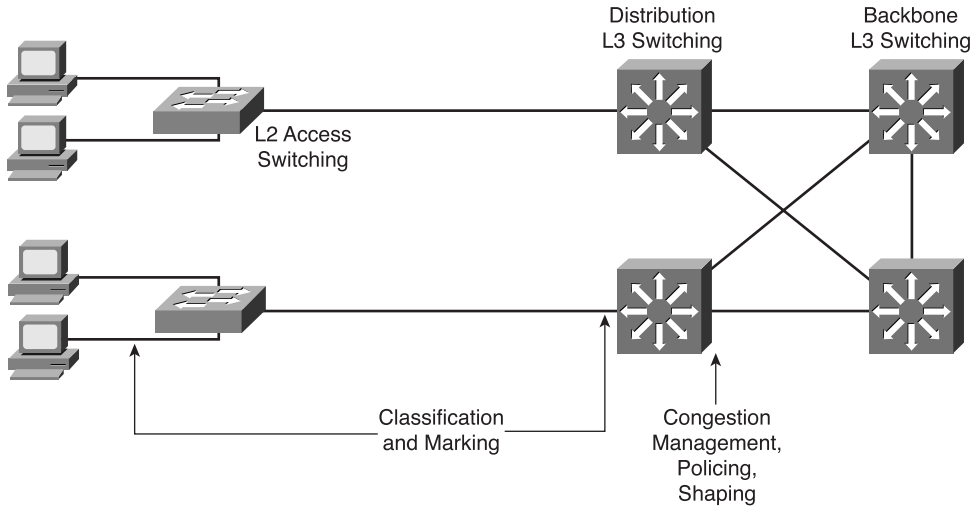
- **Congestion management**—A network interface is often congested (even at high speeds, transient congestion is observed), and queuing techniques are necessary to ensure that traffic from the critical applications is forwarded appropriately. For example, real-time applications such as VoIP and stock trading might have to be forwarded with the least latency and jitter.
- **Policing and shaping**—Policing and shaping is the process of reducing a stream of data to a predetermined rate or level. Unlike traffic shaping, in which the frames can be stored in small buffers for a short period of time, policing simply drops or lowers the priority of the frame that is out of profile.

QoS in LAN Switches

When configuring QoS features, select the specific network traffic, prioritize it according to its relative importance, and use congestion-management techniques to provide preferential treatment. Implementing QoS in the network makes network performance more predictable and bandwidth use more effective.

Figure 4-11 illustrates where the various categories of QoS are implemented in LAN switches.

Figure 4-11 *QoS in LAN Switches*



Because they do not have knowledge of Layer 3 or higher information, access switches provide QoS based only on the switch's input port. For example, traffic from a particular host can be defined as high-priority traffic on the uplink port. The scheduling mechanism of an access switch's output port ensures that traffic from such ports is served first. The proper marking of input traffic ensures the expected service when traffic passes through the distribution and core layer switches.

Distribution and core layer switches are typically Layer 3-aware and can provide QoS selectively—not only on a port basis, but also according to higher-layer parameters, such as IP addresses, port numbers, or even QoS bits in the IP packet. These switches make QoS classification more selective by differentiating the traffic based on the application. QoS in distribution and core switches must be provided in both directions of traffic flow. The policing for certain traffic is usually implemented on the distribution layer switches.

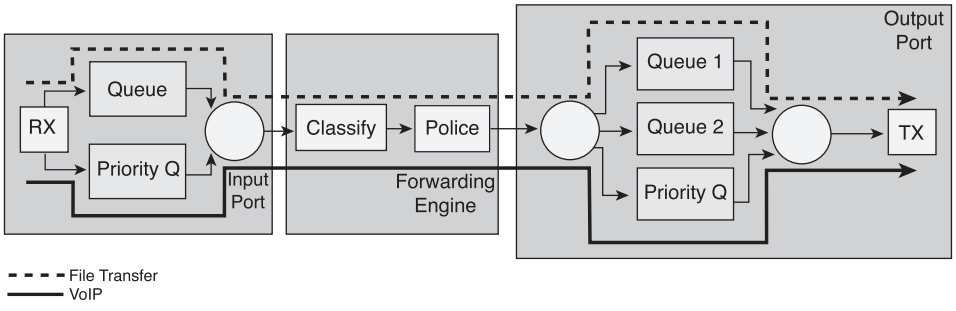
QoS Example with Voice Traffic Across a Switch

QoS for voice over IP (VoIP) consists of keeping packet loss and delay within certain tolerable levels that do not affect the voice quality. Voice requires low jitter and low packet loss. One solution would be to simply provide sufficient bandwidth at all points in the network. A better alternative is to apply a QoS mechanism at the network's oversubscribed points.

A reasonable design goal for VoIP end-to-end network delay is 150 milliseconds, a level at which the speakers do not notice the delay. A separate outbound queue for real-time voice

traffic can be provided to achieve guaranteed low delay for voice at campus speeds. Bursty data traffic, such as a file transfer, should be placed in a different queue. Packet loss is not an issue if low delay is guaranteed by providing a separate queue for voice. Figure 4-12 illustrates this situation.

Figure 4-12 *QoS for VoIP Example*



QoS maps well to the multilayer campus design. Packet classification is a multilayer service that is applied at the wiring-closet switch (access switch), which is the network's ingress point. VoIP traffic flows are recognized and then classified by their port number—the IP ToS is set to low delay voice for VoIP packets. Wherever the VoIP packets encounter congestion in the network, the local switch or router applies the appropriate congestion management based on this ToS value.

Building Access and Distribution Layers Design

In a conventional campus-wide VLAN design, network designers apply Layer 2 switching to the access layer, while the distribution layer switches support Layer 3 capabilities. In small networks, both access and distribution layers can be merged into a single switch.

Building Access Layer Considerations

The access layer aggregates the workstations or hosts on a Layer 2 device (a switch or hub). The Layer 2 node represents one logical segment and is one broadcast domain. VLAN support might be required where multiple departments coexist in the same wiring closet.

The policies implemented on the access switch are based on Layer 2 information. These policies focus on and include the following features:

- Port security
- Access speeds
- Traffic classification priorities that are defined on uplink ports

When implementing the campus infrastructure's building access submodule, consider the following questions:

- How many users or host ports are currently required in the wiring closet, and how many will it require in the future? Should the switches support fixed or modular configuration?
- What cabling is currently available in the wiring closet, and what cabling options exist for uplink connectivity?
- What Layer 2 performance does the node need?
- What level of redundancy is needed?
- What is the required link capacity to the distribution layer switches?
- How will the VLANs and STP deployed? Will there be a single VLAN, or several VLANs per access switch? Will the VLANs on the switch be unique or spread across multiple switches? The latter design was common a few years ago, but today campus-wide (or access layer-wide) VLANs are not desirable.
- Are additional features, such as port security, multicast traffic management, and QoS (traffic classification based on ports), required?

Based on the answers to these questions, the network designer can select the devices that satisfy the access layer's customer requirements. The access layer should maintain the simplicity of traditional LAN switching, with the support of basic network intelligent services and business applications.

Redundant paths can be used for failover and load balancing. Layer 2 switches can support features that are able to accelerate STP timers and provide for faster convergence and switchover of traffic to the redundant link, including BackboneFast, UplinkFast, and RSTP.

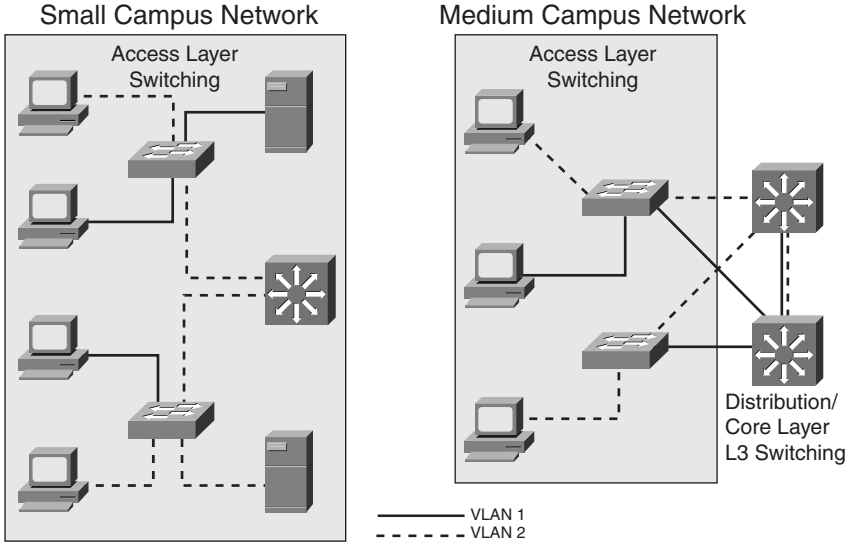
Disabling STP on a device is not encouraged because of possible loops. STP should only be disabled on carefully selected ports, and typically only on those in which the hosts are connected. Several other methods enable loop-resistant deployment of the STP, including BPDU Filtering and the BPDU guard feature on the access links where PortFast is configured. On the uplinks, BPDU skew detection, STP loop guard, and UDLD are additional measures against STP loops.

The "Layer 2 and Layer 3 Switching Design Considerations" section of this chapter discusses these STP features.

Building Access Design Examples

Figure 4-13 illustrates examples of a small and a medium-sized campus network design.

Figure 4-13 *Small and Medium Campus Access Layer Designs*



Following are some characteristics of a small campus network design:

- Network servers and workstations in small campus networks connect to the same wiring closet.
- Switches in small campus networks do not usually require high-end performance.
- The network designer does not have to physically divide the network into a modular structure (building access and building distribution modules).
- Low-end multilayer switches could provide the Layer 3 services closer to the end user when there are multiple VLANs at the access layer.
- Small networks often merge the distribution and access layers.

Because of their performance requirements, medium-size campus networks are built on Layer 2 access switches and are connected by uplinks to the distribution Layer 3 switches. This forms a clear structure of building access and building distribution modules. If redundancy is required, an additional Layer 3 switch can be attached to the network's aggregation point with full link redundancy.

Building Distribution Layer Considerations

The building distribution layer aggregates the access layer and uses a combination of Layer 2 and Layer 3 switching to segment workgroups and isolate segments from failures and broadcast storms. This layer implements many policies based on access lists and QoS settings. The distribution layer can protect the core network segment from any impact of access layer problems by implementing all the policies.

One most frequently asked question regarding implementation of a building's distribution layer is whether a Layer 2 switch is sufficient, or a Layer 3 switch must be deployed. To make this decision, answer the following questions:

- How many users will the distribution switch handle?
- What type and level of redundancy are required?
- As intelligent network services are introduced, can the network continue to deliver high performance for all its applications, such as Video On Demand, IP multicast, or IP telephony?

The network designer must pay special attention to the following network characteristics:

- **Performance**—Distribution switches should provide wire-speed performance on all ports. This feature is important because of access layer aggregation on one side and high-speed connectivity of the core module on the other side. Future expansions with additional ports or modules can result in an overloaded switch if it is not selected properly.
- **Intelligent network services**—Distribution switches should not only support fast Layer 2 and/or Layer 3 switching, but should also incorporate intelligent network services such as high availability, QoS, security, and policy enforcement.
- **Manageability and scalability**—Expanding and/or reconfiguring distribution layer devices must be easy and efficient. These devices must support the required management features.

With the correct selection of distribution layer switches, the network designer can easily add new building access modules.

NOTE Layer 3 switches are usually preferred for the distribution layer switches because this layer must support intelligent network services, such as QoS and traffic filtering.

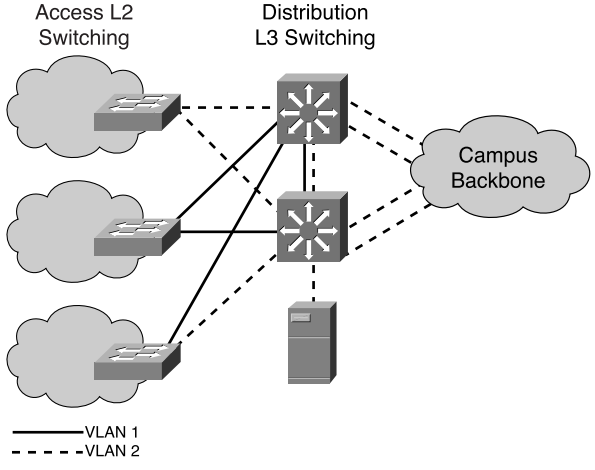
The network designer must also decide where redundancy should be implemented and which mechanisms should be used: Layer 2 and STP, or Layer 3 (routing protocol) redundant paths. If advanced STP features such as RSTP, UplinkFast, or BackboneFast are not implemented, Layer 2 redundancy and STP configuration can take up to 50 seconds. If the aforementioned features are supported and enabled on the switch, the switchover time

could be from 1 second (in the case of RSTP deployment) to 30 seconds. Routing protocols usually switch over in a few seconds (EIGRP in a redundant configuration is usually faster than OSPF because of the default 5-second shortest path first [SPF] algorithm delay recalculation).

Building Distribution Layer Example

Figure 4-14 illustrates a sample network. In this figure, each access layer module has two equal-cost paths to the distribution module switches. The distribution layer switch also has two equal-cost paths to the backbone to ensure fast failure recovery and possibly load sharing.

Figure 4-14 *Redundancy in the Building Distribution Layer*



Because of the redundancy in this example, the network designer must also address STP on the distribution layer, particularly when the possibility of Layer 2 loops exists. Figure 4-14 illustrates the access layer that is connected to both distribution switches, which are also directly interconnected. If the same VLAN spreads across all links, STP must be implemented on the access and distribution switches. If the link to the access layer fails, STP recovery time might also be a concern. STP features such as UplinkFast, BackboneFast, or RSTP can reduce the time taken to switch from one active link to another. If the connectivity to the campus backbone is based on Layer 3, STP on those ports is not necessary.

Campus Backbone Design

Low price per port and high-port density can govern wiring closet environments, but high-performance wire-rate multilayer switching drives the campus backbone design.

A campus backbone should be deployed where three or more buildings are to be connected in the enterprise campus. Backbone switches reduce the number of connections between the distribution layer switches and simplify the integration of enterprise campus modules (such as the Server Farm and Edge Distribution modules). Campus backbone switches are Layer 2 and Layer 3 switches that are primarily focused on wire-speed forwarding on all interfaces. Backbone switches are differentiated by the level of performance achieved per port rather than by high port densities.

When implementing the campus backbone, the first issue the network designer must solve is the switching mechanism—and consequently, the entire campus backbone design (Layer 2, Layer 3, or mixed Layer 2/Layer 3). Other issues to consider include the following:

- The Layer 2/Layer 3 performances needed in the campus network's backbone.
- The number of high-capacity ports for distribution layer aggregation and connection to other campus modules, such as the Server Farm or Distribution Edge.
- Redundancy requirements. To provide adequate redundancy, at least two separate switches (ideally located in different buildings) must be deployed.

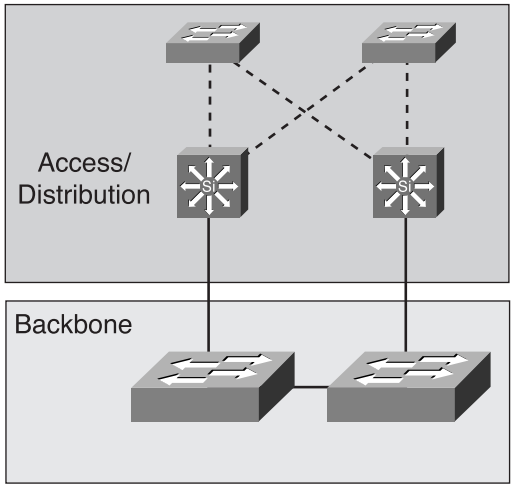
The following sections discuss different campus backbone designs.

Layer 2 Campus Backbone Design

The simplest Layer 2-based backbone consists of a single Layer 2 switch that represents a single VLAN, with a star topology toward distribution layer switches. A single IP subnet is used in the backbone, and each distribution switch routes traffic across the backbone subnet. In this case, no loops exist, STP does not put any links in blocking mode, and STP convergence does not affect the backbone.

Figure 4-15 illustrates another Layer 2 campus backbone that has two switches for backbone redundancy and a single VLAN configured per switch.

Figure 4-15 *Single VLAN Layer 2 Campus Backbone Design*



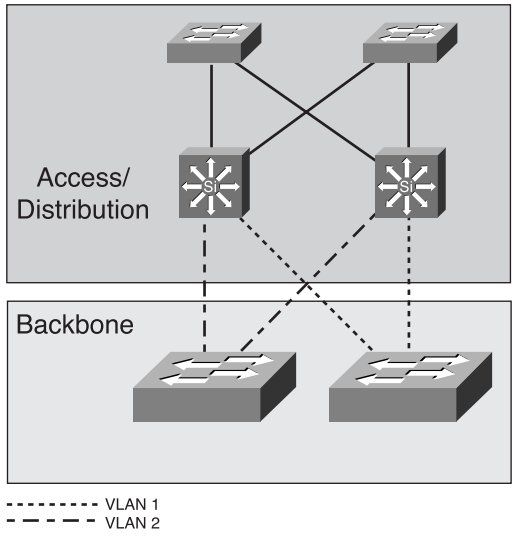
To prevent STP loops in the design in Figure 4-15, the distribution switch links to the backbone must be defined as routed interfaces (this is possible because the distribution switches are Layer 3 switches), not as VLAN trunks. This solution can lead to problems resulting from numerous Layer 3 connections between the routers that are attached to the Layer 2 backbone—especially if this includes a large number of routers.

NOTE One of the additional drawbacks of a Layer 2-switched backbone is the lack of mechanisms to efficiently handle broadcast and multicast frames—an entire backbone is a single broadcast domain. Although the broadcast/multicast suppression feature can prevent the flood of such packets, this traffic increases CPU utilization on network devices and consumes available bandwidth in the backbone network.

Split Layer 2 Campus Backbone Design

You can implement an alternative solution that uses Layer 2 in a backbone with two VLAN domains, each on one switch but without a connection between the switches. Figure 4-16 illustrates this solution, which is known as a *split Layer 2 backbone*.

Figure 4-16 *Split Layer 2 Campus Backbone Design*



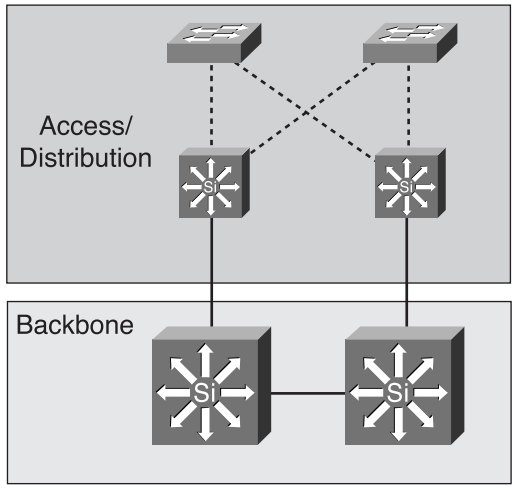
The advantage of this design is the two equal-cost paths across the backbone, which provide for fast convergence and possible load sharing.

Although the design increases high availability, it still suffers from the usual Layer 2 problems with inefficient handling of broadcast and multicast frames. In the particular case shown in Figure 4-16, the broadcast domain is limited to a single switch (that has one VLAN).

Layer 3 Campus Backbone Design

For large enterprise networks, a single or two-broadcast domain backbone is not the recommended solution. As illustrated in Figure 4-17, the most flexible and scalable campus backbone consists of Layer 3 switches.

Figure 4-17 *Layer 3 Campus Backbone Design*



Layer 3-switched campus backbones provide several improvements over the Layer 2 backbone, including the following:

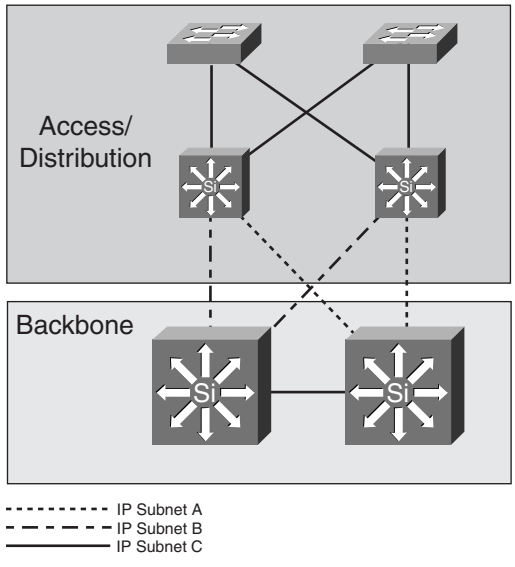
- A reduced number of connections between Layer 3 switches. Each Layer 3 distribution switch (router) connects to only one Layer 3 campus backbone switch (Layer 3). This implementation simplifies any-to-any connectivity between distribution and backbone switches.
- Flexible topology without any spanning-tree loops. There is no Layer 2 switching in the backbone or on the distribution links to the backbone because all links are routed links. Arbitrary topologies are supported because of the routing protocol used in the backbone.
- Multicast and broadcast control in the backbone.
- Scalable to an arbitrarily large size.
- Better support for intelligent network services due to Layer 3 support in the backbone switches.

One of the main considerations when using Layer 3 backbone switches is Layer 3 switching performance. Layer 3 switching requires more sophisticated devices for high-speed packet routing. Modern Layer 3 switches support routing in the hardware, even though the hardware might not support all the features. If the hardware does not support a selected feature, it must be performed in software; this can dramatically reduce the data transfer. For example, QoS and access list tables might not be processed in the hardware if they have too many entries, thereby resulting in switch performance degradation.

Dual-path Layer 3 Campus Backbone Design

As illustrated in Figure 4-18, dual links to the backbone are usually deployed from each distribution layer switch to provide redundancy and load sharing in the Layer 3-switched campus backbone.

Figure 4-18 *Dual-path Layer 3 Campus Backbone Design*



This design's main advantage is that each distribution layer switch maintains two equal-cost paths to every destination network. Thus, recovery from any link failure is fast and higher throughput in the backbone results because load sharing is possible.

The core switches should deliver high-performance, multilayer switching solutions for an enterprise campus. They should also address requirements for the following:

- Gigabit speeds
- Data and voice integration
- LAN/WAN/MAN convergence
- Scalability
- High availability
- Intelligent multilayer switching in backbone/distribution and server aggregation environments

NOTE In some situations, the campus backbone can be implemented as a mixture of Layer 2 and Layer 3 designs. Special requirements, such as the need for auxiliary VLANs for VoIP traffic and private VLANs for Server Farms, can influence the design decision.

The auxiliary VLAN feature allows IP phones to be placed into their own VLAN without any end-user intervention.

On the other hand, the private VLAN feature simplifies Server Farm designs in which the servers are separated and have no need to communicate between themselves (such as in hosting services implementation). These servers can be placed in a private VLAN with proper port assignments on the switches to ensure that servers do not communicate between themselves, while at the same time maintaining communication with external world. (More information on private VLANs is available in *Securing Networks with Private VLANs and VLAN Access Control Lists*, at www.cisco.com/en/US/products/hw/switches/ps700/products_tech_note09186a008013565f.shtml.)

Network Management Module Integration

Another consideration associated with the campus backbone is the question of network management module integration. Although a campus-wide management VLAN was used in the past, this approach has been replaced by the Layer 3 switching approach, in which the Network Management module is on its own subnet and its traffic is routed across the network.

Server Placement

Within a campus network, servers may be placed locally in a building access module, a building distribution module, or a separate Server Farm module. Servers also have numerous physical connectivity options. This section discusses these topics.

Local Server in a Building Access Module

If a server is local to a certain workgroup that corresponds to one VLAN and all workgroup members and the server are attached to an access layer switch, most of the traffic to the server is local to the workgroup. This scenario follows the conventional 80/20 rule for campus traffic distribution. If required, an access list at the distribution module switch could hide these servers from the enterprise.

Server in a Building Distribution Module

In some mid-size networks, a network designer can also attach servers to distribution switches. The designer can define these servers as building-level servers that communicate

with clients in different VLANs but that are still within the same physical building. A network designer can create a direct Layer 2-switched path between a server and clients in a VLAN in two ways:

- With multiple network interface cards (NICs), making a direct attachment to each VLAN.
- With a trunk connection or a separate VLAN on the distribution switch for the common servers.

If required, the network designer can selectively hide servers from the rest of the enterprise by using an access list on the distribution layer switch.

Server Farm

Centralizing servers in an enterprise campus is a common practice. In some cases, the enterprise consolidates services into a single server. In other cases, servers are grouped at a data center for physical security and easier administration. These centralized servers are grouped into a Server Farm module.

Server Directly Attached to Backbone

The campus backbone generally transports traffic quickly, without any limitations. Servers in medium-sized networks can be connected directly to backbone switches, making the servers only one hop away from the users. However, the need for additional traffic control in the backbone arises out of the need for controlled server access. Policy-based (QoS and ACL) control for accessing the Server Farm is implemented in the Building Distribution or Edge Distribution modules.

Switches in the Server Farm Module

Larger enterprises place common servers in a Server Farm module and connect them to the backbone via multilayer distribution switches. Because of high traffic load, the servers are usually Fast Ethernet-attached, Fast EtherChannel-attached, or even Gigabit Ethernet-attached. Access lists at the Server Farm module's Layer 3 distribution switches implement the controlled access to these servers. Redundant distribution switches in a Server Farm module and solutions such as the Hot Standby Router Protocol (HSRP) provide fast failover. (Chapter 3 discusses HSRP.) The Server Farm module distribution switches also keep all server-to-server traffic off the backbone.

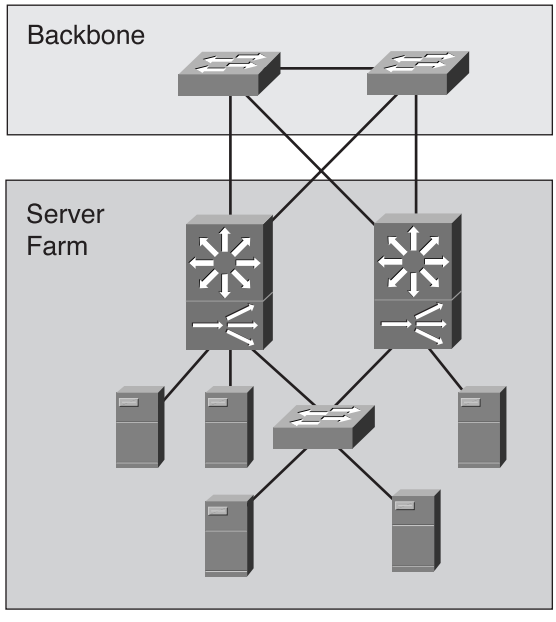
Rather than being installed on only one server, modern applications are distributed among several servers. This approach improves application availability and responsiveness. Therefore, placing servers in a common group (in the Server Farm module) and using intelligent Layer 3 switches provides the applications and servers with the required scalability, availability, responsiveness, throughput, and security.

Server Farm Design Guidelines

As shown in Figure 4-19, you can implement the Server Farm as a high-capacity building block attached to the campus backbone by using a modular design approach. One of the main concerns regarding the Server Farm is that it receives the majority of the traffic from the entire campus. Random frame drops can result because the uplink ports on switches are frequently oversubscribed. To guarantee that no random frame drops exist for business-critical applications, the network designer must apply QoS mechanisms to the server links.

NOTE Switch oversubscription occurs when some switches allow more ports (bandwidth) in the chassis than the switch's hardware is capable of transferring through its internal structure.

Figure 4-19 *Server Farm Design*



You must design the Server Farm switches with less oversubscription than switches that reside in the building access or distribution modules have. For example, if the campus consists of a few distribution modules that are connected to the backbone with Fast Ethernet, you should attach the Server Farm module to the backbone with either Gigabit Ethernet or multiple Fast Ethernet links.

The switch performance and the bandwidth of the link from the Server Farm to the backbone are not the only considerations. You must also evaluate the server's capabilities. Although server manufacturers support a variety of NIC connection rates (such as Gigabit Ethernet), the underlying network operating system might not be able to transmit at maximum capacity. As such, oversubscription ratios can be raised, thereby reducing the Server Farm's overall cost.

Server Connectivity Options

Servers can be connected in several different ways. For example, a server can attach by one or two Fast Ethernet connections. If the server is dual-attached, one interface can be active while the other is in hot standby. Installing multiple single-port or multiport NICs in the servers allows dual-homing using various modes, resulting in higher server availability.

Within the Server Farm, multiple VLANs can be used to create multiple policy domains as required. If one particular server has a unique access policy, a network designer can create a unique VLAN and subnet for that server. If a group of servers has a common access policy, the entire group can be placed in a common VLAN and subnet. Access control lists can be applied on the interfaces of the Layer 3 switches.

NOTE Several other solutions improve server responsiveness and evenly distribute the load to them. Content switches provide a robust front end for Server Farms by performing functions such as load balancing of user requests across Server Farms to achieve optimal performance, scalability, and content availability. See Chapter 3 for more information on content switching.

The Effect of Applications on Switch Performance

Server Farm design requires that you consider the average frequency at which packets are generated and the packets' average size. These parameters are based on the enterprise applications' traffic patterns and number of users of the applications.

Interactive applications, such as conferencing, tend to generate high packet rates with small packet sizes. In terms of application bandwidth, the packets-per-second (pps) limitation of the Layer 3 switches might be more critical than the throughput. Applications that involve large movements of data, such as file repositories, transmit a high percentage of full-length packets. For these applications, uplink bandwidth and oversubscription ratios become key factors in the overall design. Actual switching capacities and bandwidths vary based on the mix of applications.

Designing Connectivity to the Remainder of the Enterprise Network

The Enterprise Campus functional area's Edge Distribution module connects the Enterprise Campus with the Enterprise Edge functional area.

Recall that the Enterprise Edge functional area is comprised of the following four modules:

- **E-commerce module**—Enables enterprises to successfully deploy e-commerce applications.
- **Internet Connectivity module**—Provides internal users with connectivity to Internet services.
- **Virtual Private Network (VPN)/Remote Access module**—Terminates VPN traffic that is forwarded by the Internet Connectivity module, remote users, and remote sites. This module also terminates dial-in connections that are received through the Public Switched Telephone Network (PSTN).
- **WAN module**—Uses different WAN technologies for routing the traffic between remote sites and the central site.

The Edge Distribution module filters and routes traffic into the core (the campus backbone). Layer 3 switches are the key devices that aggregate edge connectivity and provide advanced services. The switching speed is not as important as security in the Edge Distribution module, which isolates and controls access to servers that are located in the Enterprise Edge modules (for example, servers in an E-commerce module or public servers in an Internet Connectivity module). These servers are closer to the external users and therefore introduce a higher risk to the internal campus. To protect the core from threats, the switches in the Edge Distribution module must protect the campus against the following attacks:

- **Unauthorized access**—All connections from the Edge Distribution module that pass through the campus backbone must be verified against the user and the user's rights. Filtering mechanisms must provide granular control over specific edge subnets and their ability to reach areas within the campus.
- **IP spoofing**—IP spoofing is a hacker technique for impersonating another user's identity by using their IP address. Denial of Service (DoS) attacks use the IP spoofing technique to generate server requests using the stolen IP address as a source. The server does not respond to the original source, but it does respond to the stolen IP address. DoS attacks are a problem because they are difficult to detect and defend against; attackers can use a valid internal IP address for the source address of IP packets that produce the attack. A significant amount of this type of traffic renders the attacked server unavailable and interrupts business.
- **Network reconnaissance**—Network reconnaissance (or discovery) sends packets into the network and collects responses from the network devices. These responses provide basic information about the internal network topology. Network intruders use this approach to find out about network devices and the services that run on them.

Therefore, filtering traffic from network reconnaissance mechanisms before it enters the enterprise network can be crucial. Traffic that is not essential must be limited to prevent a hacker from performing network reconnaissance.

- **Packet sniffers**—Packet sniffers, or devices that monitor and capture the traffic in the network, represent another threat. Packets belonging to the same broadcast domain are vulnerable to capture by packet sniffers, especially if the packets are broadcast or multicast. Because most of the traffic to and from the Edge Distribution module is business critical, corporations cannot afford this type of security lapse. Layer 3 switches can prevent such an occurrence.

NOTE Chapter 3 and Chapter 9, "Evaluating Security Solutions for the Network," further discuss security threats.

With the correct selection of network edge switches, all connectivity and security requirements can be met. The basic request, such as the need for ACLs, requires a switch that is Layer 3-aware. Only switches that provide such advanced features as intrusion detection can satisfy the requirements for tighter restrictions.

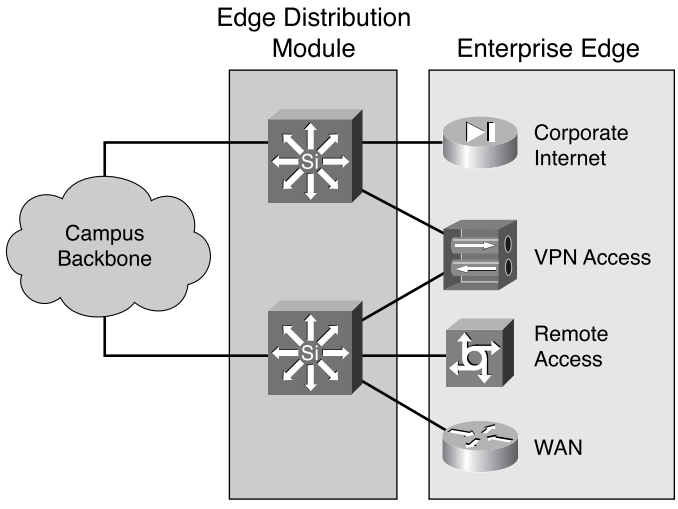
Design Guidelines for the Edge Distribution Module

Figure 4-20 illustrates an example of Edge Distribution design. In terms of overall functionality, the campus Edge Distribution module is similar to the Campus Building Distribution submodule in some respects. Although both modules use access control to filter traffic, the Edge Distribution module can rely on Enterprise Edge modules to perform additional security functions to some degree. Both modules use Layer 3 switching to achieve high performance, but the Edge Distribution module can offer additional security functions because its performance requirements are not as high. The Edge Distribution module provides the last line of defense for all traffic that is destined for the Campus Infrastructure module. This line of defense includes mitigation of spoofed packets, mitigation of erroneous routing updates, and provisions for network layer access control.

Alternatively, the Edge Distribution module can be combined with the Campus Backbone submodule if performance requirements are not as stringent; this is similar to combining the Server Farm module and Campus Building Distribution submodule.

Security can be implemented in this scenario by using intrusion detection line cards in the Layer 3 switches. (Network Intrusion Detection Systems [NIDSs] reduce the need for external appliances at the points where the critical edge modules connect to the campus; performance reasons can dictate that dedicated intrusion detection is implemented in the various edge modules, as opposed to simply the Edge Distribution module.)

Figure 4-20 *Edge Distribution Design Example*



Summary

This chapter discussed campus network design fundamentals using a multilayer design and the positioning of switches in campus modules.

Geography, application requirements, data link layer technology, cabling, and type of traffic forwarding are the factors you must consider when designing a campus network.

Location of nodes and the distance between them determine a campus network's geography. Intra-building, inter-building, and distance remote building are geographical structures that serve as guides to determine Enterprise campus requirements.

Characterization of applications that are used on a network can determine enterprise traffic patterns. Four types of application communication are client-client, client-distributed server, client-Server Farm, and client-Enterprise Edge.

Switched technology has many benefits over shared technology, including higher bandwidth support, larger network diameter, additional Layer 2 and Layer 3 services, and high availability.

Deciding whether to use Layer 2 or Layer 3 switching involves the consideration of network service capabilities, the size of the network segments, and maximum network failure convergence time that can be tolerated.

The most common physical transmission media used in modern networks are twisted-pair cables (copper) and optical cables (fiber). The choice of physical media depends on bandwidth and the distance between devices.

Network traffic affects the campus design. Considerations include application traffic patterns, the presence of multicast traffic, and the presence of delay-sensitive traffic. Multicast design considerations can prevent flooding of the traffic to all switched ports.

The Building Access module aggregates the workstations or hosts on a Layer 2 device. The distribution layer aggregates the access layer and uses a combination of Layer 2 and Layer 3 switching to segment workgroups and isolate segments from failures and broadcast storms.

The Campus Backbone and Server Farm modules require fast and resilient connectivity. Campus backbone switches are Layer 2 and Layer 3 switches that are primarily focused on wire-speed forwarding on all interfaces.

In the Edge Distribution model, the speed of switching is not as important.

References

For additional information, refer to the following resources:

- Introduction to Gigabit Ethernet, www.cisco.com/warp/public/cc/techno/media/lan/gig/tech/gigbt_tc.htm
- Gigabit Campus Network Design—Principles and Architecture, www.cisco.com/warp/public/cc/so/neso/lnso/cpso/gcnd_wp.htm
- Gigabit Networking Gigabit Ethernet Solutions, www.cisco.com/warp/partner/synchronicd/cc/techno/lnty/etty/ggetty/tech/gesol_wp.htm

NOTE You must be a registered user to access this document.

- SAFE: A Security Blueprint for Enterprise Networks, www.cisco.com/go/safe
- LAN Design Guide for the Midmarket, www.cisco.com/warp/public/cc/pd/si/casi/ca3500xl/prodlit/lan_dg.htm

NOTE Appendix B, “References,” lists all the Web sites referenced in this chapter.

Case Study and Simulation Exercise

This case study is a continuation of the DJMP Industries case study we introduced in Chapter 2, "Applying Design Principles in Network Deployment."

Key Point: Case Study General Instructions

Use the scenarios, information, and parameters provided at each task of the ongoing case study. If you encounter ambiguities, make reasonable assumptions and proceed. For all tasks, use the initial customer scenario and build on the solutions provided thus far.

You can use any and all documentation, books, white papers, and so on.

In each task, you act as a network design consultant. Make creative proposals to accomplish the customer's business needs. Justify your ideas when they differ from the provided solutions.

Use any design strategies and internetworking technologies you feel are appropriate.

The final goal for each case study is a paper solution; you are not required to provide the specific product names.

Appendix G, "Answers to Review Questions, Case Studies, and Simulation Exercises," provides a solution for each task based on assumptions made. There is no claim that the provided solution is the best or only solution. Your solution might be more appropriate for the assumptions you made. The provided solution helps you understand the author's reasoning and offers a way for you to compare and contrast your solution.

Case Study: Enterprise Campus Design

Complete these steps:

Step 1 You might want to review the DJMP Industries Case Study Scenario in Chapter 2.

Step 2 Propose the optimal campus design that addresses the scenario requirements (switched solution, redundancy, servers in a separate segment, and so on).

Simulation 1: Shared Versus Switched LAN

This exercise is a paper-only version of the simulation that the simulation tool actually performed and includes the results it provided. Review the scenario and simulation results and answer the questions.

Scenario

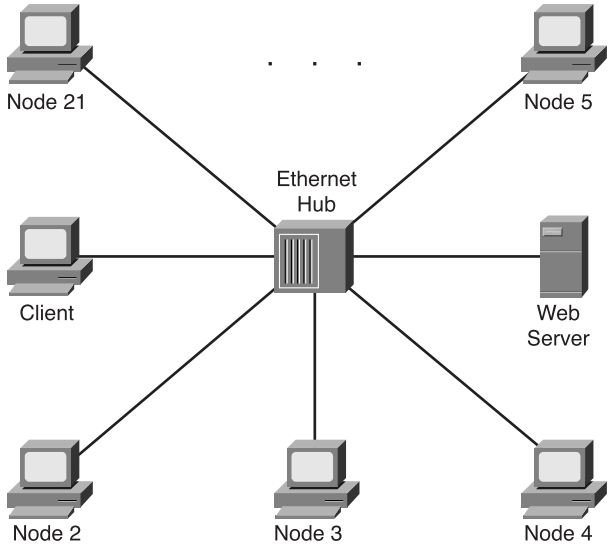
The customer (DJMP Industries) plans to restructure its flat campus network, which consists of workstations and servers that are located in the central building and building A. The company is considering Ethernet switching technology as a replacement for the 10BaseT Ethernet hubs. You have been asked to determine what effect the introduction of the switches might have on the load of the links and to estimate the network's responsiveness and utilization with respect to the existing applications.

To provide some proof of future network efficiency, you will model FTP and HTTP performance on the network using shared and then switched Ethernet platforms.

Client Accessing Server in Unloaded Shared Ethernet

The customer has provided the information about its existing network and the number of users. As illustrated in Figure 4-21, you began the initial network behavior evaluation by simulating the load on the LAN links, which was posed by a single client accessing the web server.

Figure 4-21 *Single Client Accessing Web Server on Unloaded Shared Ethernet*

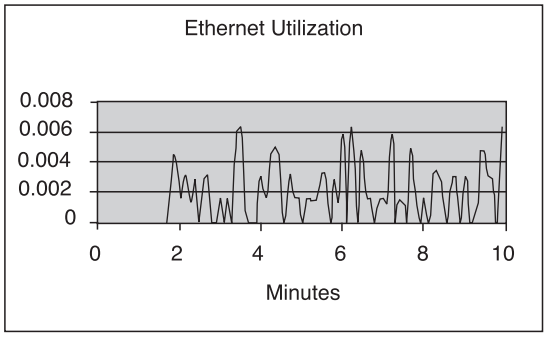


You performed the simulation (using 10-minute intervals), observed the effect of traffic growth, and compared the results among different scenarios.

The relevant statistics of interest for this case are the link (Ethernet) utilization and the HTTP response times.

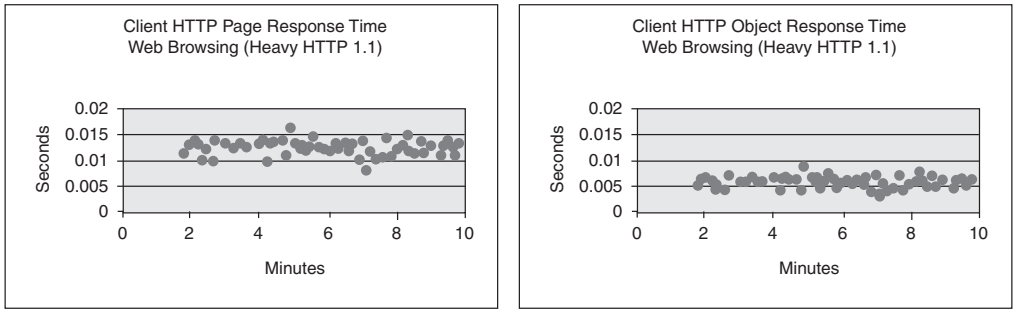
The graph in Figure 4-22 shows the network load's simulation results that resulted from the HTTP session between the client and the server. The low Ethernet utilization number indicates that the HTTP traffic exchanged between the client and the server does not represent a significant load in the network.

Figure 4-22 *Ethernet Utilization on Unloaded Shared Ethernet*



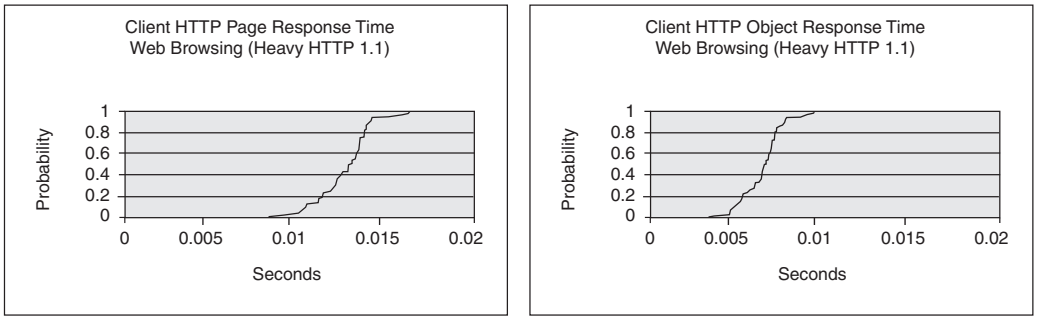
The graphs in Figure 4-23 show the simulation results of the HTTP response times. On average, the HTTP page response times are within the range of 0.01 and 0.015 seconds, whereas the HTTP object response times vary from approximately 0.004 to 0.01 seconds (every HTTP page consists of several objects).

Figure 4-23 *HTTP Response Times on Unloaded Shared Ethernet*



The graphs in Figure 4-24 show the simulation results of the probability that the HTTP response time is equal to a particular value.

Figure 4-24 *Probability of HTTP Response Times on Unloaded Shared Ethernett*

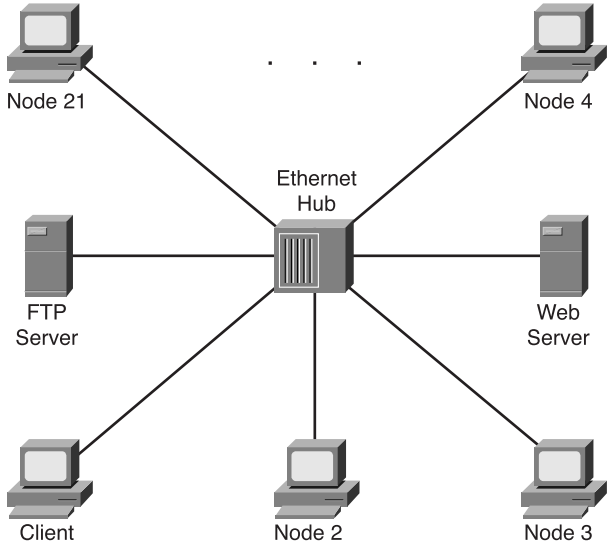


1 What can you observe from the graphs in Figures 4-23 and 4-24?

Client Accessing Server in Loaded Shared Ethernet

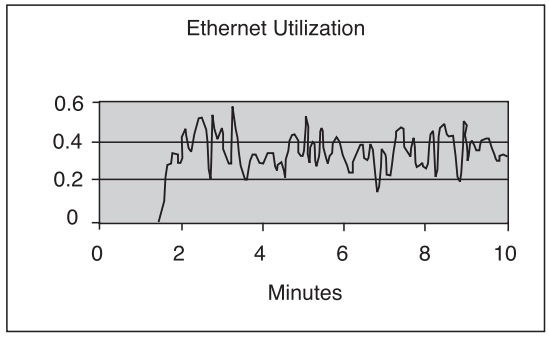
Your task now is to create a scenario in which the background traffic is simulated to provide a more realistic picture of the ongoing traffic in the network. The client continues to access the web server while all the other clients concurrently initiate FTP sessions to an FTP server; as illustrated in Figure 4-25, a separate FTP server is introduced to eliminate the effect of the server utilization. Therefore, the HTTP session is tested in a heavily-loaded, shared Ethernet network.

Figure 4-25 *Single Client Accessing the Web Server on Loaded Shared Ethernet*



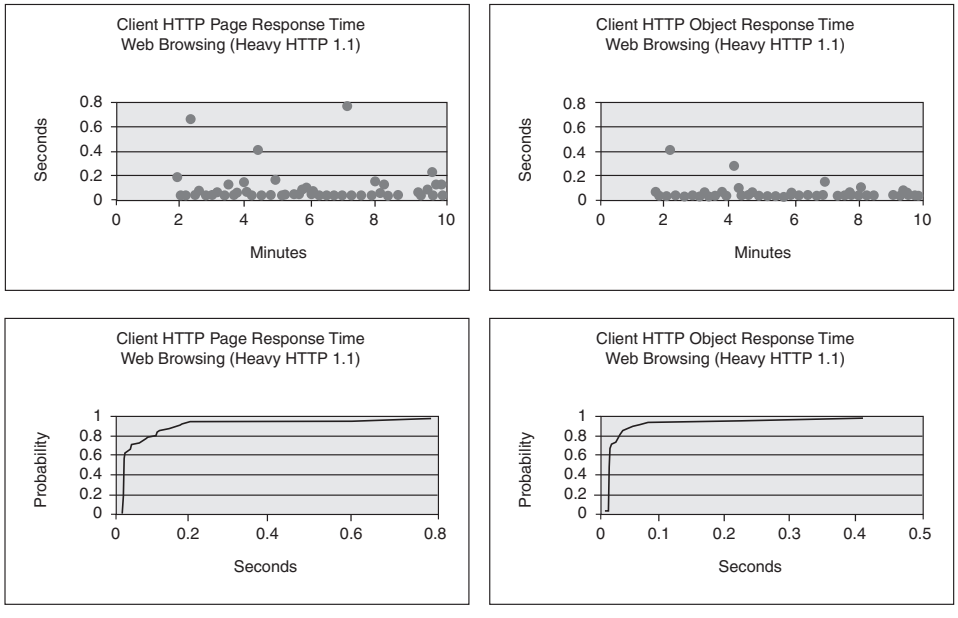
You performed the simulation and compared the results with those from the previous simulation. The graph in Figure 4-26 describes the increased network utilization as a result of the concurrent FTP and HTTP conversations.

Figure 4-26 *Ethernet Utilization on a Loaded Shared Network*



The next step is to observe the HTTP response times again. When examining the graphs in Figure 4-27, you notice that, in general, the results match those that were obtained in the unloaded network. There are some deviations, presumably because of the retransmissions that lower the probability of an immediate response. The delayed responses seem evenly distributed throughout the observed interval.

Figure 4-27 *HTTP Response Times on Loaded Shared Ethernet*



2 What can you determine from the results? What is the reason for the delayed HTTP responses?

Introducing Switched Ethernet

In the third simulation scenario, the shared Ethernet is replaced with switched Ethernet, which Figure 4-28 shows being implemented with a single LAN switch. The traffic pattern remains the same as in the previous scenario—the client is accessing a web server while all other clients are accessing an FTP server.

Figure 4-28 *Single Client Accessing the Web Server on Switched Loaded Ethernet*

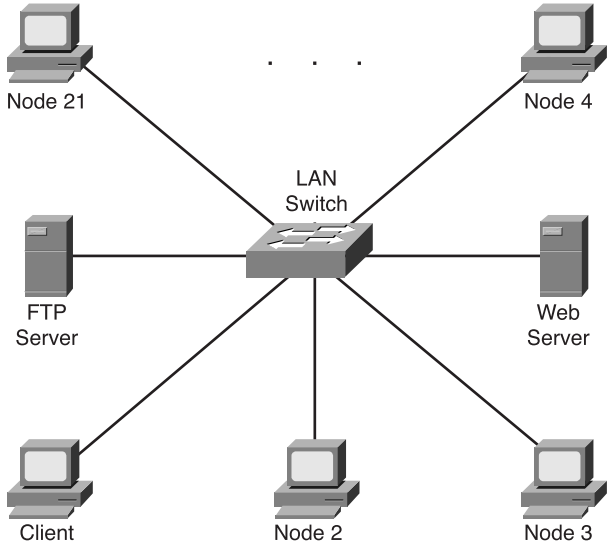
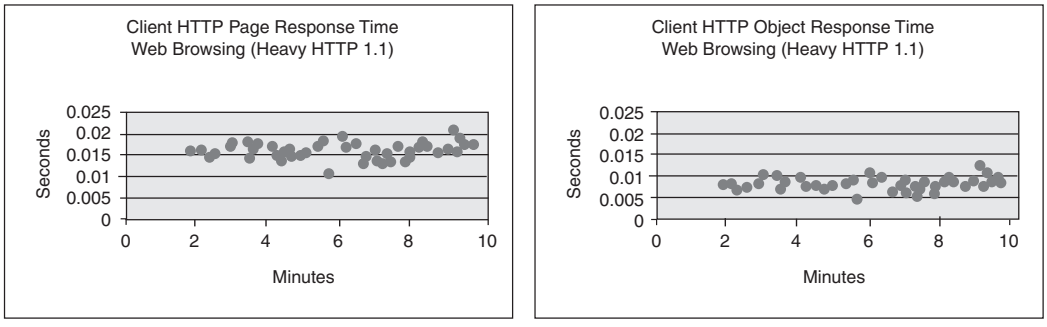


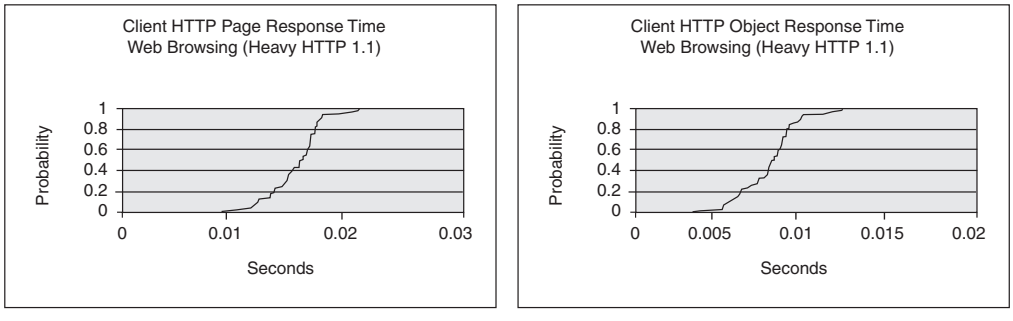
Figure 4-29 shows the results of this simulation. By examining the HTTP response time carefully, it seems that the background FTP traffic does not significantly affect the web communication. Everything is back to normal, the HTTP response times are constantly low, and there is no sign of individual deviations that could compromise the overall statistic numbers.

Figure 4-29 *HTTP Response Times on Loaded Switched Ethernet*



The graph in Figure 4-30 illustrates the probability of receiving a prompt HTTP response. The possibility is almost as high as when a stand-alone HTTP session was simulated (with no background traffic). This leads you to the conclusion that switching technology might be the obvious solution.

Figure 4-30 *HTTP Response Probabilities on Loaded Switched Ethernet*



3 You concluded that the introduction of the Layer 2 switch represents a significant improvement in this case. How did you determine this from the previous graphs?

Simulation 2: Layer 2 Versus Layer 3 Switching

This exercise is a paper-only version of the simulation that the simulation tool actually performed, including the results the tool provided. Review the scenario and the simulation results and answer the questions.

Scenario

This simulation inspects the impact of Layer 2 versus Layer 3 switching on the load in various parts of the structured campus network.

After successfully deploying the switching technology, the company is considering further improvements to its campus network design. It has already finished some baseline wiring work in the central building and in Building A and is facing some Layer 2 and Layer 3 design issues.

You decided to model the company's network to match the existing situation using the following architecture:

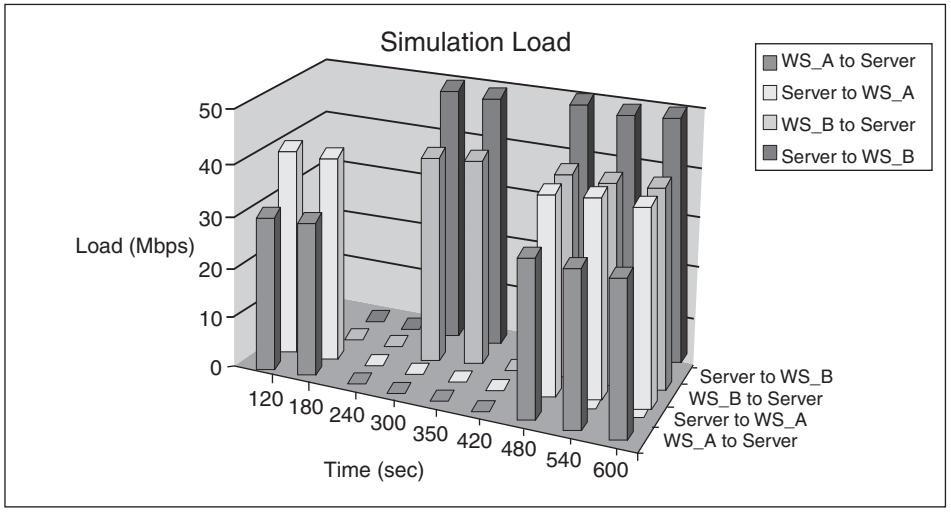
- Each building contains distribution-layer switches, to which the access-layer (wiring closet or data center concentrator) switches are connected.
- The distribution layer devices are connected via two central core switches (the campus backbone).
- The whole campus is fully redundant.

To provide comparable results, you need a reference traffic flow. Therefore, you decided to focus solely on the communication between the two workstations—WS_A and WS_B—that are located in different floors of building A, and the server in the central building.

Initial Traffic

In the simulation, Workstations A and B communicate with the server using various loads, as illustrated by the graph in Figure 4-31.

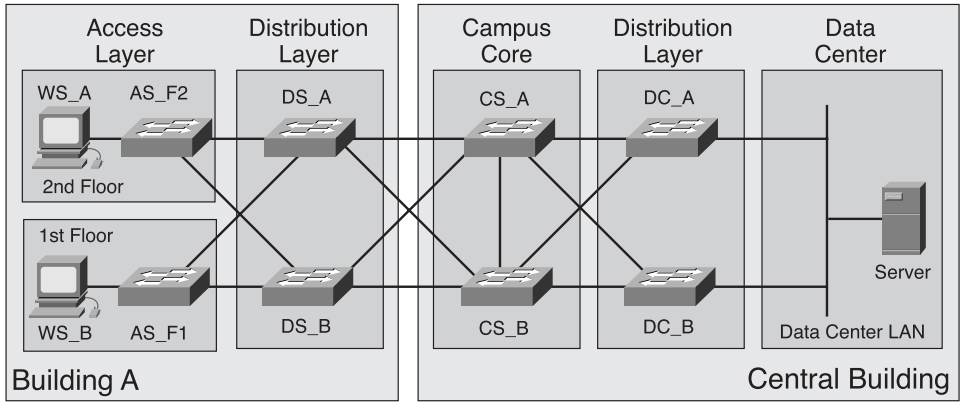
Figure 4-31 *Simulation Load*



Layer 2 Only Design

As shown in Figure 4-32, you began the simulation by turning on the Layer 2 functionality on all switches in the campus network. Soon you realized that, even in the highly redundant Layer 2 network, the number of possible paths reduces to only one, as determined by STP. STP computes loop-free networks, and any redundant links belonging to the same LAN or VLAN are placed in the blocking state and cannot be used.

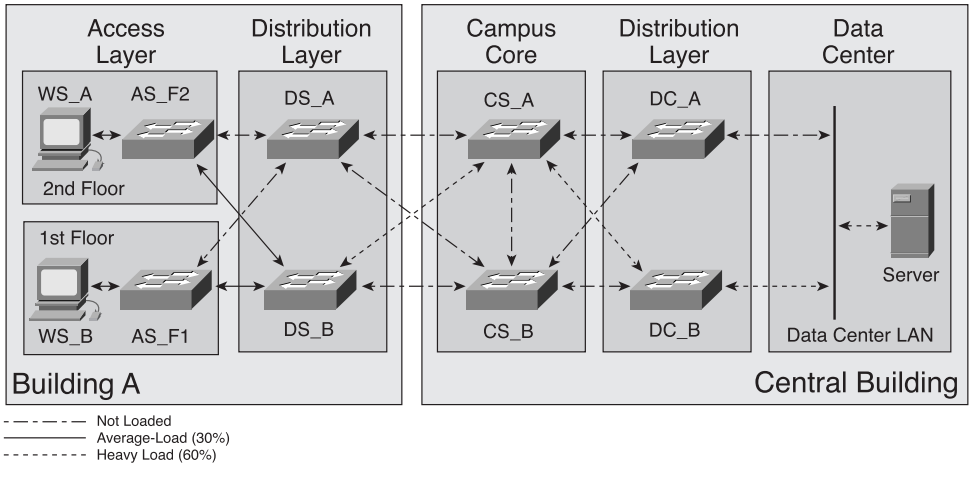
Figure 4-32 *Layer 2 Only Design*



Loaded Network

Figure 4-33 depicts the result of simulating 10 minutes of traffic originated by both workstations toward the server, and vice versa. The average-loaded links (30 percent) appear as solid lines, and the heavily-loaded links (60 percent) appear as dotted lines. The resulting dashed and dotted arrows indicate that the load is not balanced; specifically, all traffic moves over a single path: DS_B → CS_A → DC_B.

Figure 4-33 *Layer 2 Only Loaded Network*

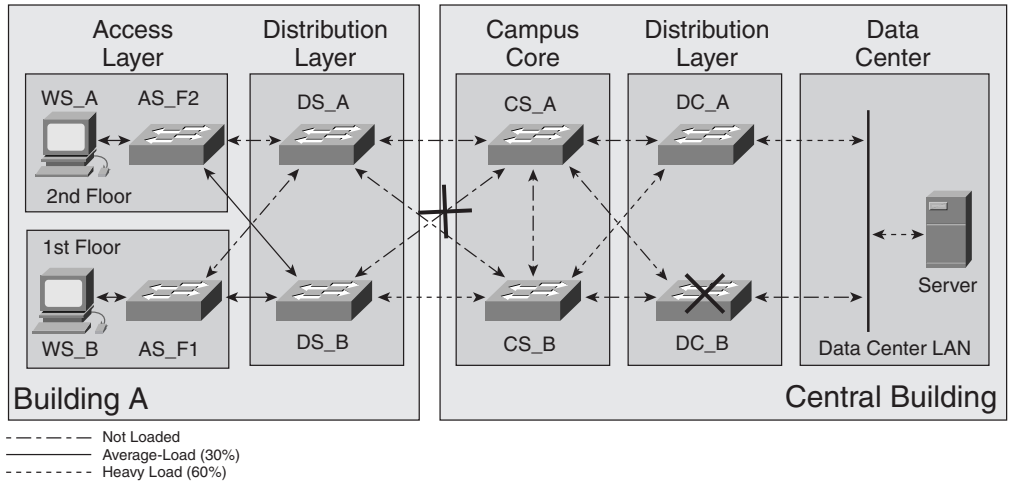


Link Failure

Use of redundant links terminating at separate devices helps increase the network's reliability. This is especially true for the observed case, in which you expect that the link or node failure would neither impact the network (at least not for a longer period) nor result in a load imbalance.

To prove this, you studied the effect of the link and node failure on the network performance by tearing down the DS_B → CS_A link and afterwards disabling the DC_B node. The resulting graph, which is illustrated in Figure 4-34, indicates that the traffic is simply redirected over the alternative path, DS_B → CS_B → DC_A.

Figure 4-34 *Link Failure on the Layer 2 Only Loaded Network*



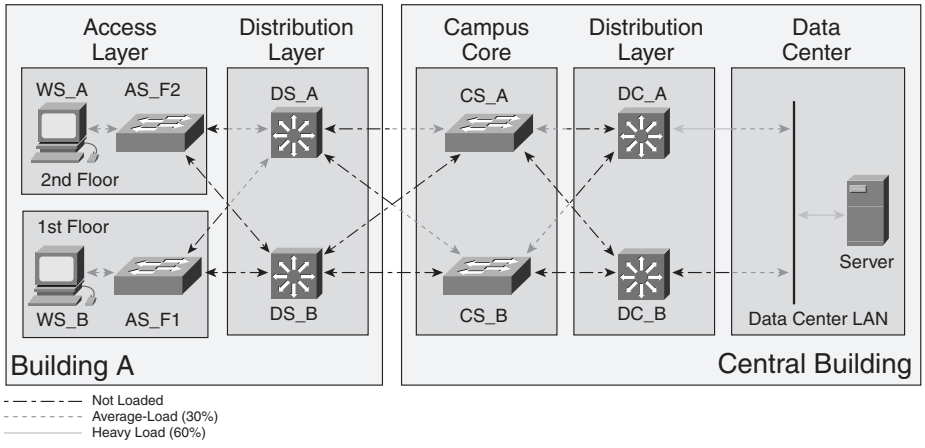
4 Does the traffic immediately start using the original path once the link or node has fully recovered?

Layer 3 Switching in Distribution

Next, you decided to replace distribution-layer Layer 2 switches with Layer 3 switches, thereby eliminating the STP path selection restrictions. This was expected to improve the efficiency of the distribution to core link usage.

Figure 4-35 presents the results of the simulation. The traffic is perfectly balanced from the ingress Layer 3 switch all the way to the destination. The sharing is proportional on pairs of source-destination distribution switches, so all the distribution switches are equally loaded (see the arrows representing the load: dotted for average load and solid for heavy load). The access layer contains the only remaining sub-optimal paths.

Figure 4-35 *Balanced Traffic with Layer 3 Switching in the Distribution Layer*



5 Examining the results in Figure 4-36, you might notice that no load sharing occurs in building A's access layer. Is this a result of the default routing on the workstations using distribution switch DS_A for the primary exit point, or a result of the attached Layer 2 switch placing the secondary port in the blocking mode?

Traffic Flow

The graph in Figure 4-36 shows the path (using thick lines) taken by the packet that is originated by workstation WS_A and destined for the server in the central building. It is obvious that the network resources are used more fairly.

Figure 4-36 *Traffic Flow from WS_A to the Server*

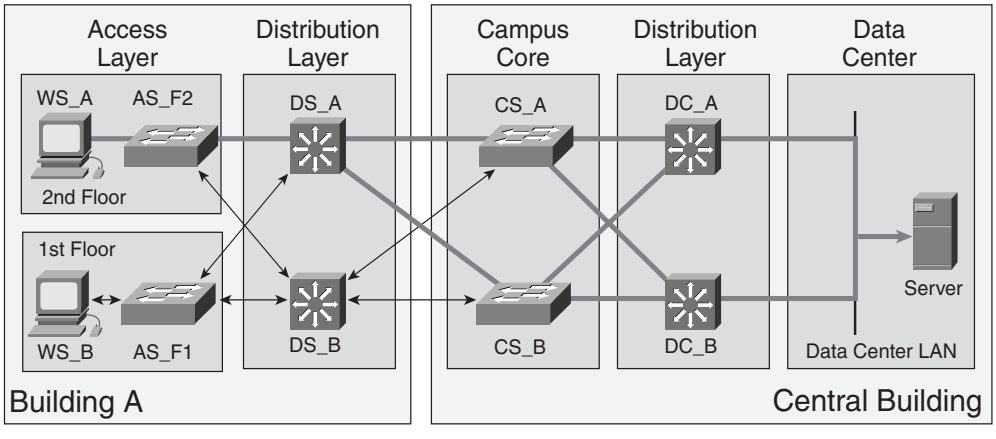
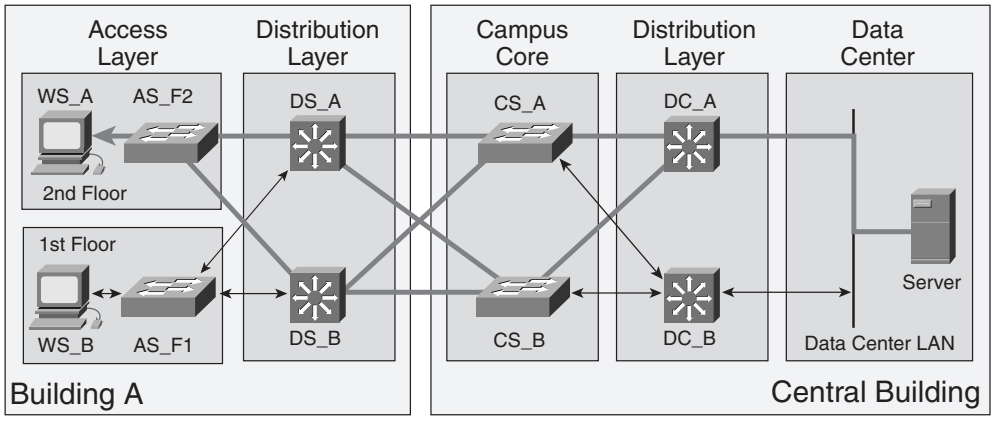


Figure 4-37 shows the path (using thick lines) the packets take in the opposite direction, from the server toward the workstation WS_A. The server uses default routing to send the packets out of the local LAN and therefore does not utilize the redundant path at all.

Figure 4-37 *Traffic Flow from Server to the WS_A*



Failure Resilience

The network is now tested against severe failure events, such as link loss, by simulating a failure of the CS_A → DC_A link. Figure 4-38 shows the result; the traffic from WS_A to the server is represented by a thick line. As expected, the network does not change its behavior under link failure. The load balancing from the ingress Layer 3 switch to the destination is still perfect, but on a reduced topology. The load distribution ratio on DS_A → CS_A versus DS_A → CS_B is 1:2 because the load is shared between distribution-layer next hops.

Figure 4-38 *Link Failure Scenario Showing WS_A to Server Traffic*

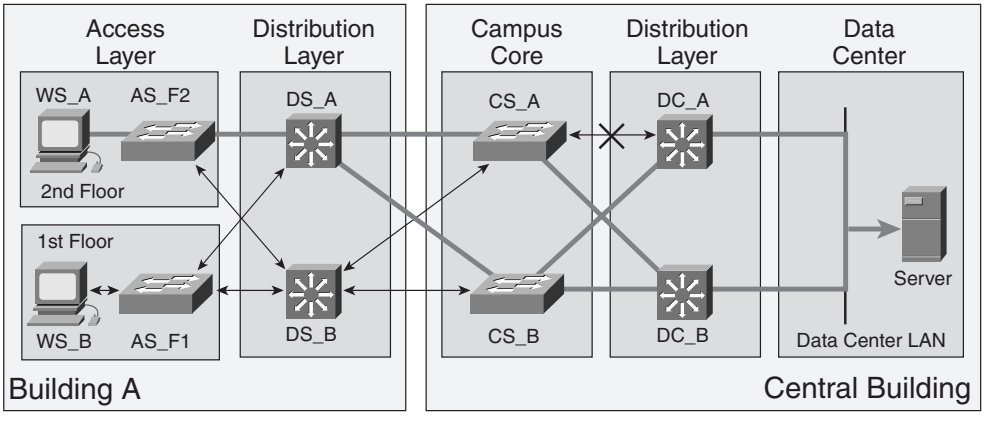
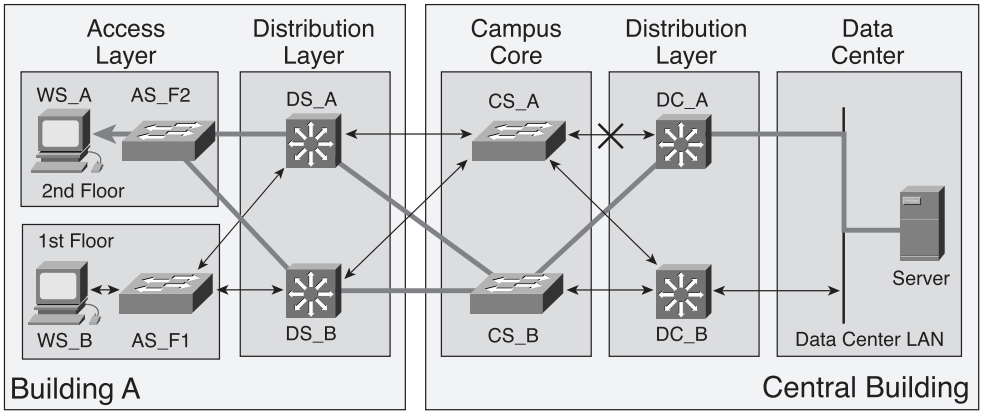


Figure 4-39 illustrates the return path from the server to WS_A (shown as thick lines).

Figure 4-39 *Link Failure Scenario Showing Server to WS_A Traffic*

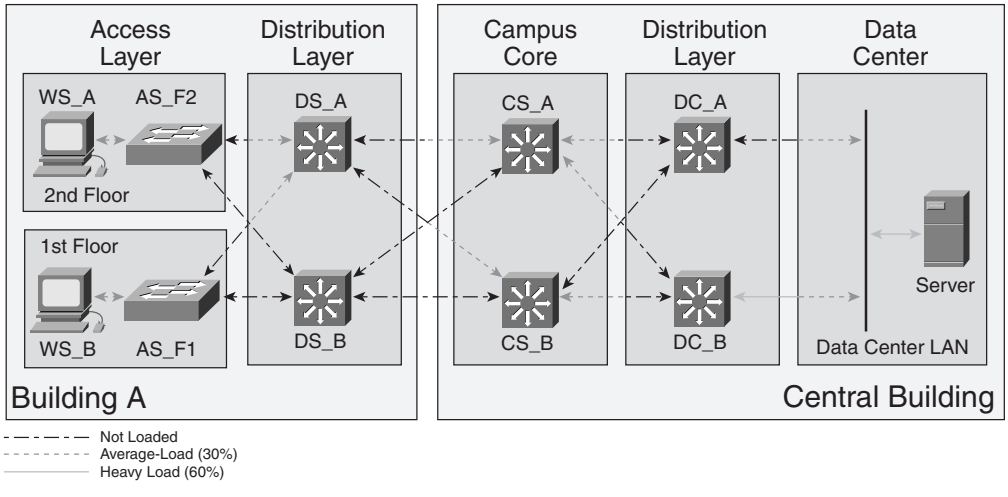


6 In Figure 4-39, why is the return path completely bypassing the CS_A switch?

Layer 3 Switching in Core and Distribution

At this point, you change the core so that the core and distribution layer switches are all Layer 3 switches. As illustrated in Figure 4-40, the simulated load is perfectly shared from the distribution layer across the core on a hop-by-hop basis.

Figure 4-40 *Layer 3 Switching Results in a Balanced Load*

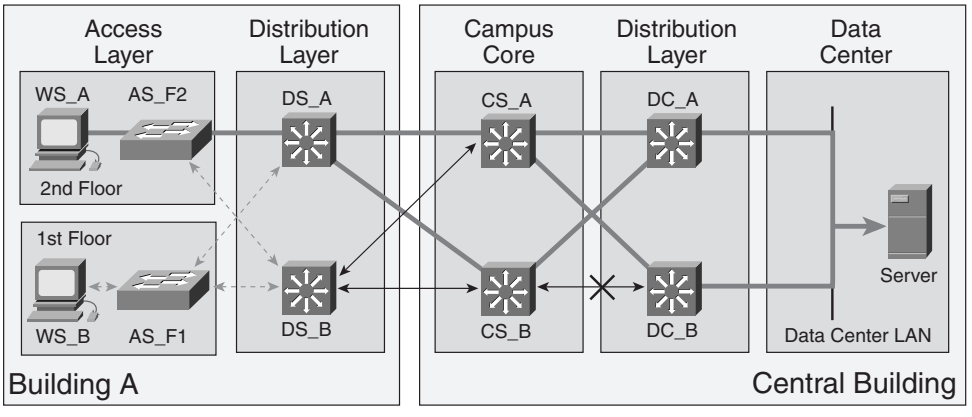


Load Sharing Under Failure

Next you simulated failure of the link CS_B to DC_B. Figure 4-41 illustrates the resulting path (shown as thick lines) taken by the WS_A traffic to the server. The way the load sharing is done is comparable to the previous case, with the distribution Layer 3 switches and Layer 2 switching in the core.

NOTE The actual impact of Layer 3 switches in the core can only be seen if the convergence after the failure is taken into account.

Figure 4-41 *Link Failure Is Accommodated by WS_A to Server Traffic with the Layer 3 Core*



7 What is the load distribution ratio on DS_A → CS_A versus the DS_A → CS_B link in Figure 4-41? Explain.

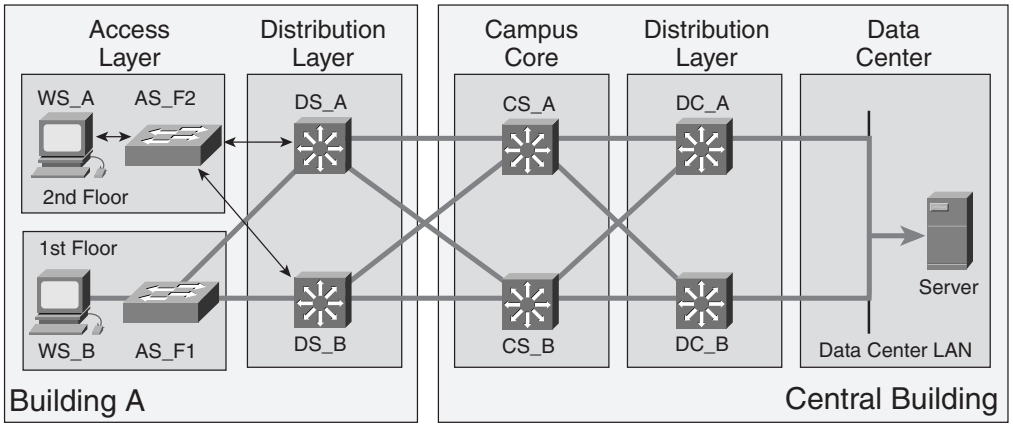
Layer 3 Access Switch

No load sharing occurs in the access-layer LAN or VLAN if the access-layer switch is a Layer 2 switch and all the workstations use the same default gateway (distribution-layer switch). To achieve load sharing in the access layer, the workstations must be configured to use different next hops (DS_A and DS_B, in this case) for their default routes.

In this scenario, the AS_F1 access-layer switch is upgraded to a Layer 3 switch to achieve more optimal load sharing in the access layer.

Figure 4-42 illustrates the result of the simulation (shown as thick lines): load sharing from AS_F1 toward DS_A and DS_B is perfect.

Figure 4-42 *Load Sharing in Access the Layer with a Layer 3 Switch*



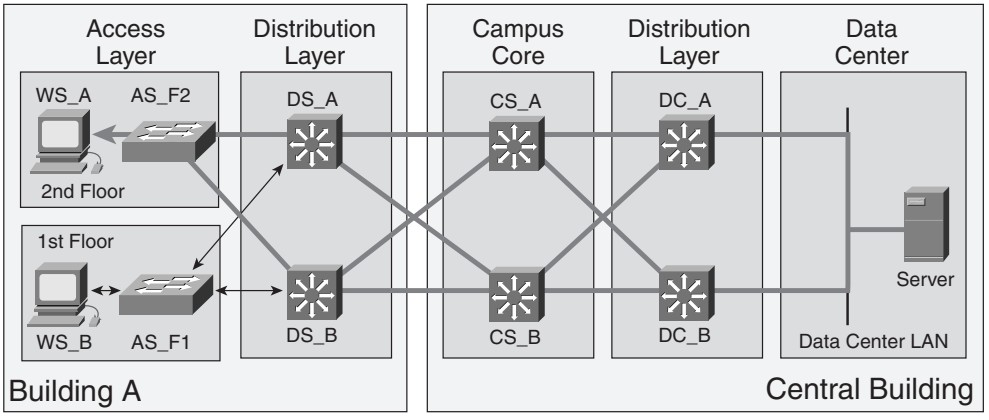
8 The workstation WS_B is not running any routing protocol; rather, it depends on the default routing. What is a proper next-hop address?

IP Routing Process on the Server

In the last scenario, OSPF is configured on the server. The server starts participating in the campus routing and can rely on OSPF to load-share its traffic toward the workstations.

Figure 4-43 shows the result of the server to WS_A path simulation (shown as thick lines): the load distribution is achieved from the access layer to the destination.

Figure 4-43 *Load Sharing with the Server Running OSPF*



9 Running a routing protocol is one way to force the server to forward packets to both distribution-layer switches. Can you think of any other option?

Review Questions

Answer the following questions and then refer to Appendix G, "Answers to Review Questions, Case Studies, and Simulation Exercises," for the answers.

1 What factors must be considered when designing a campus network?

2 A company keeps all its servers and workstations within one building. What geographical design structure should be chosen?

3 What are some of the differences between inter-building and distant remote network geographical structures?

4 The Engineering department has requirements to use a bandwidth- and processor-intensive research application. What type of application communication should it use?

5 What are some of the benefits of using LAN switches over hubs?

6 Some users in a department use an application that generates many broadcast frames, which results in up to 10 Mbps bandwidth utilization. Which of the following solutions is the most optimal in this case?

   a Provide 100 Mbps or higher connections to all users in a domain

   b Limit the number of broadcast frames in a domain for all department users

   c Optimize the application

   d Put the application users into a separate broadcast domain

7 What type of cable would you recommend for connecting two switches that are 115 m apart?

8 What is the intended result of the application characterization process?

9 With default Spanning Tree Protocol parameters, how long could it take before the redundant link is available when the currently active link fails?

10 Compare the range and bandwidth specifications of UTP, MM fiber, and SM fiber.

11 What are the benefits of using Layer 3 switches over Layer 2 switches?

12 The users in an organization are divided along their workgroup lines into VLANs. Workgroup servers are located within these workgroup VLANs. The organization has also placed mail and web servers, to which all corporate users have access, in a separate VLAN. What is the expected traffic flow in this organization?

13 A company is using video on demand, which utilizes IP multicast as part of its distance-learning program. The routers are configured for IP multicast. Taking into account that the majority of the LAN switches are Layer 2 switches, which protocol should be enabled on the LAN switches to reduce flooding?

14 Which Enterprise Campus modules typically have both high availability and high performance requirements?

15 What is the difference between the 80/20 rule and the 20/80 rule?

16 A link between the building distribution and campus backbone is oversubscribed, yet carries mission-critical data along with Internet traffic. How would you ensure that the mission-critical applications are not adversely affected by the bandwidth limitations?

17 What are two uses of redundant paths?

18 A corporate network is spread over four floors. There is a Layer 2 switch on each floor, each with more than one VLAN. One connection from each floor leads to the basement, where all WAN connections are terminated and all servers are located. Traffic between VLANs is essential. What type of device should be used in the basement?

19 What applications might require the network to handle multicast traffic?

20 What functions does the Building Distribution submodule provide?

21 What is the main focus of the campus backbone?

22 An organization requires a highly available core network and uses IP telephony for all of its voice communication, both internal and external. Which devices and topology would you recommend for the campus backbone design?

23 A company has mission-critical applications hosted on common servers that are accessible to selected employees throughout the company's multiple buildings. Where and how would you recommend that these servers be placed within the network?

24 What is the function of the Edge Distribution module?