Designing Campus LANs

Design Implementation Guide

Bill Kelly
kelly@cisco.com
Director, Technical Marketing
Cisco Systems Inc.

Terri Quinn-Andry
tquinn@cisco.com
Product Marketing Engineer
Cisco Systems Inc.

Introduction

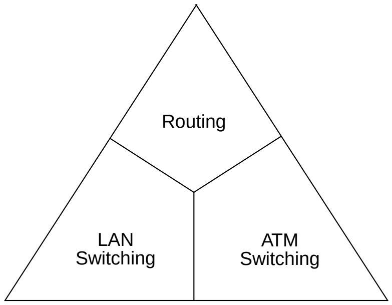
With the advent of technologies like LAN switching, virtual LANs (VLANs), and layer 3 switching, building campus LANs today has become significantly more complex than in the past. Cisco's history is in both routing and layer 3 switching, but we have made significant investments in both ATM switching and Ethernet switching as well. We are uniquely able to provide our customers with a balance of technologies.

Three technologies are required today to build successful campus networks:

- Routing Technologies
  - — *Routing*—Can be either layer 3 switching or more traditional routing with layer 3 switching features. This technology is key to connecting LANs in a campus network. Layer 3 switching will be discussed later in this paper
- LAN Switching Technologies
  - — *Ethernet Switching*—Provides layer 2 switching and offers broadcast domain segmentation using VLANs. This is the base fabric of the network. Layer 2 switching will be covered later in this paper
  - — *Token Ring Switching*—Same as Ethernet Switching but uses Token Ring technology. Can be used as either a transparent bridge or as a source-route bridge.
  - — *FDDI Switching*—Provides switching technology in the backbone with benefits of FDDI. Useful for planned migration to Asynchronous Transfer Mode (ATM).
- ATM Switching Technologies
  - — *ATM Switching*—Offers high-speed switching technology for voice, video, and data. Operation is similar to LAN switching technologies for data operations; however, it offers high bandwidth capacity.

These technologies are summarized in Figure 1.

Figure 1. Networking Technologies.



It is useful to recognize some important trends that exist today. In the past, when purchasing a technology for enterprise networks, there were few options—either hubs or routers. Design mistakes were rare with these two technologies; hubs were for wiring closets, and routers were for the data center or main telecom operations.

With the advent of emerging technologies like ATM switching and Ethernet/Token Ring switching, today's campus LAN designs are created by purchasing separate equipment types (routers, Ethernet switches, and ATM switches) and linking them together. While individual purchase decisions may seem harmless, designers must remember that the entire network forms an enterprise network. Today it is possible to separate the technologies and build thoughtful designs using each new technology, but with no consideration for the overall integration of the network. The result is networks that have a much higher risk of network outages, downtime, and congestion than ever before.

The goal of this paper is to help create a context for the new technologies that will lead to an enterprise view of them and how they interact. The hoped-for result is the creation of networks that are more robust, cleaner, and administratively easier to build and manage than current networks.

This paper will cover the following:

- End-station networking issues overview—includes addressing issues, services offered to the end stations, and how the end stations link to routers in a network. Different protocols will be highlighted.
- Switching overview—details the differences between layer 2 and layer 3 switching functions and how each switching layer helps to solve the three main networking problems: protocols, media, and transport issues.
- Campus design issues—discusses the effects of broadcasts on the network, multimedia design issues, campus mobility, VLAN design advantages/disadvantages, VLAN strategies, VLAN technology issues, and backbone redundancy.
- Designing campus LANs—offers possible design solutions for campus LAN networks with respect to the size of the network (ranging from a very small number of users to very large networks) and with respect to the campus design issues discussed in the previous section.

In this paper the terms "router" and "gateway" are interchangeable, as are "MAC layer" and "data link layer."

End-Station Networking Issues Overview

It is important to establish a common knowledge base in regard to end-station operation in the network. Key to networking decisions are how end stations are addressed at layers 2 and 3 of the Open System Interconnection (OSI) model, how services are delivered to or acquired by end stations, and how end stations link to routers.

Layer 2 and Layer 3 Addressing

Switching uses layer 2 addressing, and routing uses layer 3 addressing.

Layer 2 addresses are determined by the manufacturer of the data communications equipment used. They are unique addresses that are derived in two parts: the Manufacturing (MFG) code and the unique identifier. The MFG code is assigned to each vendor by the IEEE. The vendor assigns a unique identifier to each board it produces.

xxxxxx.xxxxxx hexadecimal

MFG Code . Unique Identifier

For example: 00000c.58ae02, commonly denoted as 0000.OC58.ae02

where 00000c is the MFG code, and

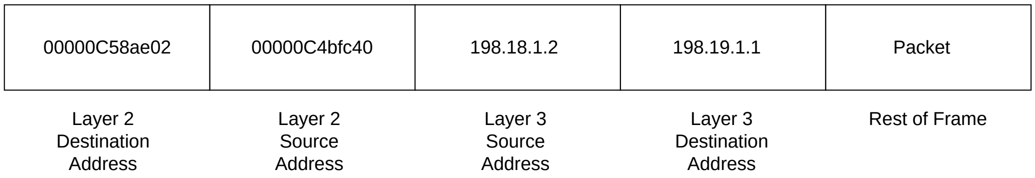
58ae02 is the unique identifier.

For the most part, layer 3 addresses are determined by the network administrator who installs a hierarchy on his or her network. Protocols such as IP, IPX, and AppleTalk use layer 3 addressing:

IP	xxx.xxx.xxx.xxx	decimal (for example, 198.18.1.2)
IPX	xxxxxxxx	hexadecimal (for example, C6120101)
AppleTalk	xxx-xx.xxxxx	decimal (for example, 257-257 257.253, where 257 is the cable range and 253 is the node address)

Figure 2 shows the layer 2 address and layer 3 address in a frame.

Figure 2. Frame Addressing



Following is a simple example to help clarify the addressing differences. Imagine if the postal service delivered mail by social security number or passport number. Each person would have a completely unique address that would always stay with them, regardless of where they lived. However, it would be impossible to always know where they were. This is a flat addressing system that is not ordered in a way that we can understand and is similar to layer 2 addressing. By introducing addresses that change if their owners change locations (that is, streets/cities/state/country), the postal service can easily deliver the mail. The house numbers along the streets give each house the uniqueness needed for successful mail delivery. This is similar to layer 3 addressing.

In networks, users have little or no control over layer 2 addressing (except in Systems Network Architecture [SNA] installations). These addresses are the equivalent of using social security numbers as network identifiers. Layer 2 addresses are fixed with a device, whereas layer 3addresses can be changed.

By creating layer 3 addresses, a network administrator creates local areas that act as single addressing units (similar to streets/cities/state/country) and assigns a number to each local entity (like a house number). In the preceding example, if people move to a new house, they change their local area addresses but their social security numbers stay the same. Thus if users move to another building, their end stations will obtain new layer 3 addresses, but their layer 2 addresses remain unchanged.

Therefore, some form of linking layer 2 and layer 3 addresses is necessary. Address linkage between layer 2 and layer 3 can occur in three ways. The first method uses a table linkage in which all end stations on a layer 2 segment participate. This method is called Address Resolution Protocol, or ARP. A table is created to link the layer 3 address with the layer 2 address:

ARP Table

Layer 2 address = Layer 3 address

AppleTalk and IP use this form of addressing. Table 1 shows an IP ARP table, and Table 2 shows an AppleTalk ARP table.

Table 1. IP ARP Table

<table>
<tr><th colspan="5">IP ARP Table</th></tr>
<tr><td colspan="5">#sho arp</td></tr>
<tr><th>Protocol Address</th><th>Age (min)</th><th>Hardware Address</th><th>Type</th><th>Interface</th></tr>
<tr><td>Internet 192.160.82.197</td><td>–</td><td>0000.0c06.417d</td><td>ARPA</td><td>Ethernet7</td></tr>
<tr><td>Internet 192.160.82.189</td><td>–</td><td>0000.0c06.411f</td><td>ARPA</td><td>Ethernet9</td></tr>
<tr><td>Internet 192.160.82.184</td><td>167</td><td>0000.0cff.c15b</td><td>ARPA</td><td>Ethernet3</td></tr>
<tr><td>Internet 192.160.82.185</td><td>22</td><td>0800.201a.f157</td><td>ARPA</td><td>Ethernet3</td></tr>
<tr><td>Internet 192.160.82.181</td><td>–</td><td>0000.0c06.4149</td><td>ARPA</td><td>Ethernet3</td></tr>
</table>

Table 2. AppleTalk ARP Table

<table>
<tr><td colspan="5">router# sho appletalk arp</td></tr>
<tr><th>Address</th><th>Age (min)</th><th>Type</th><th>Hardware Address</th><th>Encap Interface</th></tr>
<tr><td>6340.35</td><td>–</td><td>Hardware</td><td>0000.0c06.1146.0000</td><td>SNAP Fddi0</td></tr>
<tr><td>6340.182</td><td>184</td><td>Dynamic</td><td>0000.0c05.a6f4.0000</td><td>SNAP Fddi0</td></tr>
<tr><td>6341.159</td><td>169</td><td>Dynamic</td><td>0000.0c0c.224d.0000</td><td>SNAP Ethernet0</td></tr>
<tr><td>6341.215</td><td>–</td><td>Hardware</td><td>0000.0c06.1146.0000</td><td>SNAP Ethernet0</td></tr>
</table>

The second method is to include the layer 2 address as part of the layer 3 address. A typical example of this would be Novell networks in the form of:

Administrator Portion.Layer 2 address

or

<Net Number>. <MAC Address>

For example,	C0962A50.0000.0c0b.8f67
where	C0962A50 is the net number, and
	0000.0c0b.8f67 is the MAC address

The third form of linkage is used by DECnet. The layer 3 address is embedded in the layer 2 address according to an algorithm. The algorithm is shown here.

It starts with an area/node address pair. An area's value ranges from 1 to 63, and a node address can be 1 to 1023. In the algorithm, the area number is multiplied by 1024, and then the node number is added. This 16-bit decimal address is then converted to a hexadecimal number and appended to the address AA00.0400 in byte-swapped order with the least significant byte first.

For example, area/node address = 12.75

12x1024 = 12,288 + 75 = 12363

12363 decimal = 304B hex

Byte swap = 4B30

appended to AA00.0400. DECnet address = AA00.0400.4B30

Services

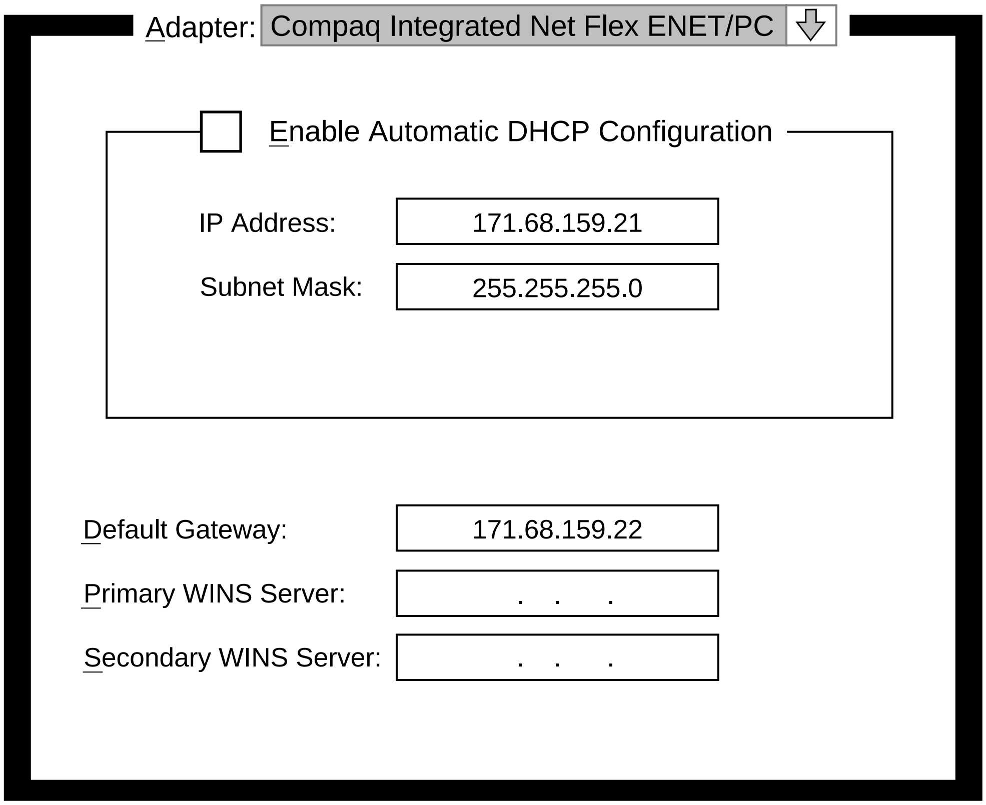
All network protocols offer services such as printing, file sharing, and imaging that are available as network resources. Each layer 3 network protocol offers different methods of advertising these services. These methods can be generalized in the following manner:

- Explicit knowledge
- Service advertisements
- Service discovery

Explicit Knowledge

Explicit knowledge is the method that IP uses to find its services. The system administrator must know explicitly which layer 3 address is the router, and which is the domain name server. Figure 3 shows an example of a PC network screen using explicit knowledge.

Figure 3. Explicit Knowledge Example



Adapter: Compaq Integrated Net Flex ENET/PC

□ Enable Automatic DHCP Configuration

IP Address: 171.68.159.21

Subnet Mask: 255.255.255.0

Default Gateway: 171.68.159.22

Primary WINS Server: . . .

Secondary WINS Server: . . .

The domain name server maps IP addresses to names so that when names are used to establish connections, the name corresponds to the IP network address. Users must use explicit names for each service. For example, to perform a file transfer using the File Transfer Protocol (FTP), a user must type a name recognized by the domain name server; for example, ftp bigserver.mycompany.com. While this might seem painful because each user must be trained to know where the servers are, it eliminates a lot of painful broadcasts. Naming conventions have also emerged so that, for example, the World Wide Web server for the entire company would be called www.mycompany.com as a convention.

While this method works for IP, it was considered too unfriendly for desktop protocols such as IPX and AppleTalk, so more automated forms of finding services were created. These methods are called service advertisement and service discovery.

Service Advertisement

Novell IPX uses service advertisements as a method of allowing end stations to find services offered to the network. In this method, each server, by default, announces in a broadcast which services it is offering every 60 seconds. This traffic can be impressive, because it takes about 75 bytes to describe each service, and some networks can have hundreds of services.

The end stations then present a list of services to either the underlying application or to the user and ask them to choose one (for example, Printer X or File Server Y).

Table 3 shows an IPX server table.

Table 3. IPX Server Table

router# sho ipx servers							
Codes: S-Static, I-Incremental, P-Periodic, H-Holddown, 5 Total IPX Servers							
Table ordering is based on routing and server info							
Type	**Name**	**Net**	**Address**	**Port**	**Route**	**Hops**	**Itf**
P	4 BERTHA	2DEB3448.0000.0000.0001:0451			3/02	2	Et0
P	4 DARKSTAR	2E1D70C2.0000.0000.0001:0451			3/02	2	Et0
P	26B CISCO_________	2E1D70C2.0000.0000.0001:0005			3/02	2	Et0
P	30C 0800092E490703C2ATHE	AB449EA1.0800.092e.4907:400C			1/01	2	Et0

One function of routers in a network is to terminate service broadcasts and maintain global service tables by protocol. Each router then maintains these tables and transmits them to other routers every 60 seconds unless the network administrator changes the default timing. An example of this process is AppleTalk zones. An AppleTalk zone is a logical group of AppleTalk networks that is not physically restricted. Routers maintain AppleTalk zone tables that are sent to the other routers at regular intervals. Table 4 shows the AppleTalk zone service table.

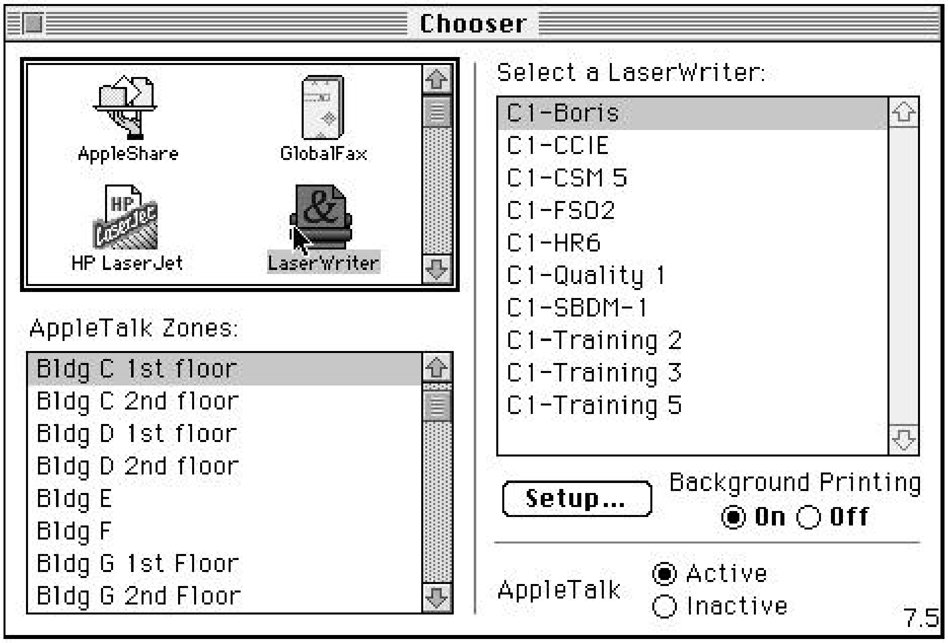
Table 4. Router Service Table

router#sho appletalk zone	
Name	**Network(s)**
OZ-Sydney	8626-8626 8637-8637 8670
Can-Montreal	42515 42512-42512 42513-42513
Europe-Dusseldorf	54536-54536 54538-54538
Europe-Madrid	24351-24351 24353
US-Hartford	7711-7711 7712-7712

Service Discovery

Service Discovery is used by the AppleTalk protocol to find services. Using this model, clients request a list of services by transmitting a broadcast searching for them. In a hierarchical layer 3 network, the logical network segments are labeled (in AppleTalk they are called network zones), and the user at an end station selects a zone. Figure 4 shows the Macintosh Chooser, which uses service discovery.

Figure 4. Service Discovery



AppleTalk Zones:

- Bldg C 1st floor
- Bldg C 2nd floor
- Bldg D 1st floor
- Bldg D 2nd floor
- Bldg E
- Bldg F
- Bldg G 1st Floor
- Bldg G 2nd Floor

Select a LaserWriter:

- C1-Boris
- C1-CCIE
- C1-CSM 5
- C1-FS02
- C1-HR6
- C1-Quality 1
- C1-SBDM-1
- C1-Training 2
- C1-Training 3
- C1-Training 5

Background Printing ● On ○ Off

AppleTalk ● Active ○ Inactive

After a zone is selected, a unicast is transmitted to the router adjacent to the zone. The router converts the unicast into a broadcast within the selected zone, and all replies are transmitted across the network to the client.

To summarize, services are obtained in three ways: explicit knowledge, service advertisement, and service discovery. Each method has its strengths and weaknesses, but all collectively affect the design of campus LANs.

Client Linkage to Routers

There are two primary methods for clients to link to routers. The first is administrative configuration, and the second is router discovery. We will discuss both methods; both have an impact on campus LAN design.

Administrative Configuration

Administrative configuration is the primary method that the IP protocol uses to link clients to their local routers. The common configuration command is gateway, which is the IP address of one of the local routers (see Figure 3). The flaw in this configuration command system is that, if the local router is turned off or fails, the end station must be reconfigured. Cisco’s Hot Standby Router Protocol (HSRP) allows administratively configured end stations to continue to operate with about a ten-second convergence, allowing two or more routers to use the same MAC address and IP network address of a virtual router. This configuration allows routers on the same network to provide backup for each other.

Router Discovery

There are three methods of router discovery:

- Route announcement
- Route listening
- Gateway discovery

Route Announcement

Route announcement involves devices actively announcing their routes. Both IP and IPX use Route Announcement. For the IP protocol, route announcement is a common and very dangerous form of router discovery. For networks that use the Routing Information Protocol (RIP), IP end stations (generally Sun and other workstations) are configured to run "**routed**," which is a command to run RIP. The danger in this configuration is that, if the network is listening to RIP from both routers and end stations, a user can misconfigure the end station with multiple IP addresses, some correct and some incorrect. If this user has inadvertently selected the address of the primary backbone, instead of its local net, then the station announces the wrong route, causing the local traffic to be routed to the wrong network.

For the Novell IPX protocol, all servers participate in route announcement (unless they are running 4.12), because each server has a unique internal IPX network number assigned to it. Novell servers routinely broadcast their route out to their network. Route announcement has one downside that might not be obvious: If a failure occurs and a backup router is configured on the same LAN (or VLAN), it takes three routing update periods (generally 30 seconds for each route update using RIP) to use the new routing path.

Route Listening

For IP, listening to routing protocols is a less common but safer method of linking hosts to routers. The same workstations that use the **routed** command can be configured to only listen to RIP. The command in UNIX is **routed -q,** where **q** is the quiet option.

The AppleTalk protocol uses route listening by selecting the first Routing Table Maintenance Protocol (RTMP) update it hears as its local gateway. A disadvantage here is that if the routers that are running AppleTalk disagree about the topology, zones, and so on of the AppleTalk network, then multiple Macintoshes on the same network could see radically different topologies.

Gateway Discovery

For IP and OSI protocols, a separate protocol exists that announces only the fact that there is a router on the local wire. End stations can take that router announcement and use it to "discover" gateways. In OSI, this protocol is called the End System-to-Intermediate System (ES-IS) Protocol, where ES is the workstation and IS is the router. For IP, Cisco created an obscure protocol called Router Discovery Protocol (RDP) that was standardized by the Internet Engineering Task Force (IETF) as the Internet Control Message Protocol (ICMP) Router Discovery Protocol (IRDP). Shareware is available from Cisco's FTP directory (ftp@cisco.com) for Sun workstations to use this protocol.

Router Linkage Summary

For IP protocols, many linkage methods exist, but the most common is to administratively configure the location of the router. This is done for PCs, not workstations, that are running **routed.** For other protocols such as IPX and AppleTalk, routing must operate on the local LANs to allow end stations to discover their paths to the network. The campus network design implications for protocols such as IPX, IP, and AppleTalk are that service information and replies, routing information, and zone information will be broadcast by each server, router, and so forth. This traffic is categorized either as useful and is negligible in terms of bandwidth, or bandwidth intensive and degrades overall network performance, depending on the overall campus design.

Switching Overview

In data communications today, all switching and routing equipment performs two basic operations:

- *Switching Data Packets*—This is generally a store-and-forward operation in which a frame arrives on an input media and is transmitted to an output media.
- *Maintenance of Switching Operations*—In this operation, switches build and maintain switching tables and search for loops, and routers build and maintain both routing tables and service tables.

There are two methods of switching data packets, layer 2 and layer 3 switching. We will discuss these methods next.

Layer 2 and Layer 3 Switching

Switching is the process of taking an incoming frame from one interface and delivering it out through another interface. The difference between layer 2 and layer 3 switching is the type of information inside the frame that is used to determine the correct output interface. Basically, layer 2 switching switches frames based on MAC address information, and layer 3 switching switches frames based on network-layer information. Figure 5 shows a layer 2 destination address, and Figure 6 shows a layer 3 destination address.

Figure 5. Layer 2 Destination Address

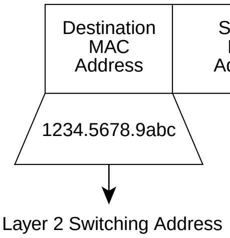
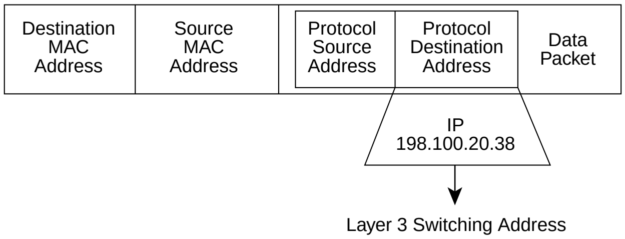


Figure 6. Layer 3 Destination Address



In this paper, a packet contains the data information and the network layer addresses. A frame contains the packet plus the MAC layer addresses.

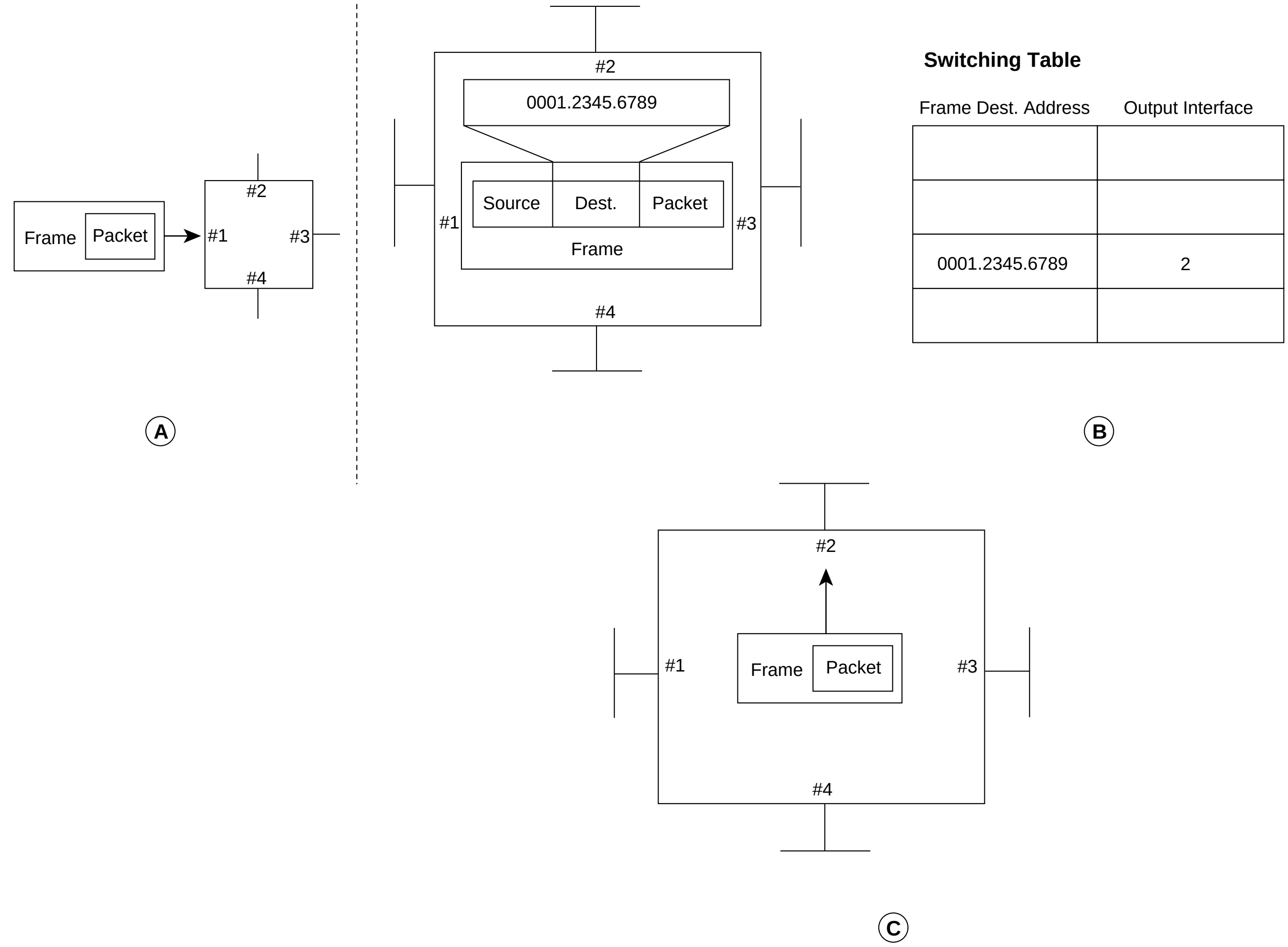
To further differentiate, layer 2 switching is performed by looking at a destination MAC address within a frame. It looks at the frame's destination address and sends it to the appropriate interface if it knows the destination address location. Layer 2 switching builds and maintains a switching table, keeping track of which MAC addresses belong to each port or interface, as shown in Table 5.

Table 5. Switching Table

#sho cam dynamic	
Destination MAC Address	**Destination Ports**
00-00-0c-06-41-7c	3/3
00-00-0c-02-b9-01	3/3
00-80-24-07-8c-60	3/2

If the layer 2 switch does not know where to send the frame, it broadcasts the frame out all its ports to the network to learn the correct destination. When the packet's reply comes back, the switch learns the location of the new address and adds the information to the switching table. Layer 2 switching does not look inside a packet for network-layer information. Figure 7 displays the layer 2 switching process for a known destination address.

Figure 7. Layer 2 Switching Process

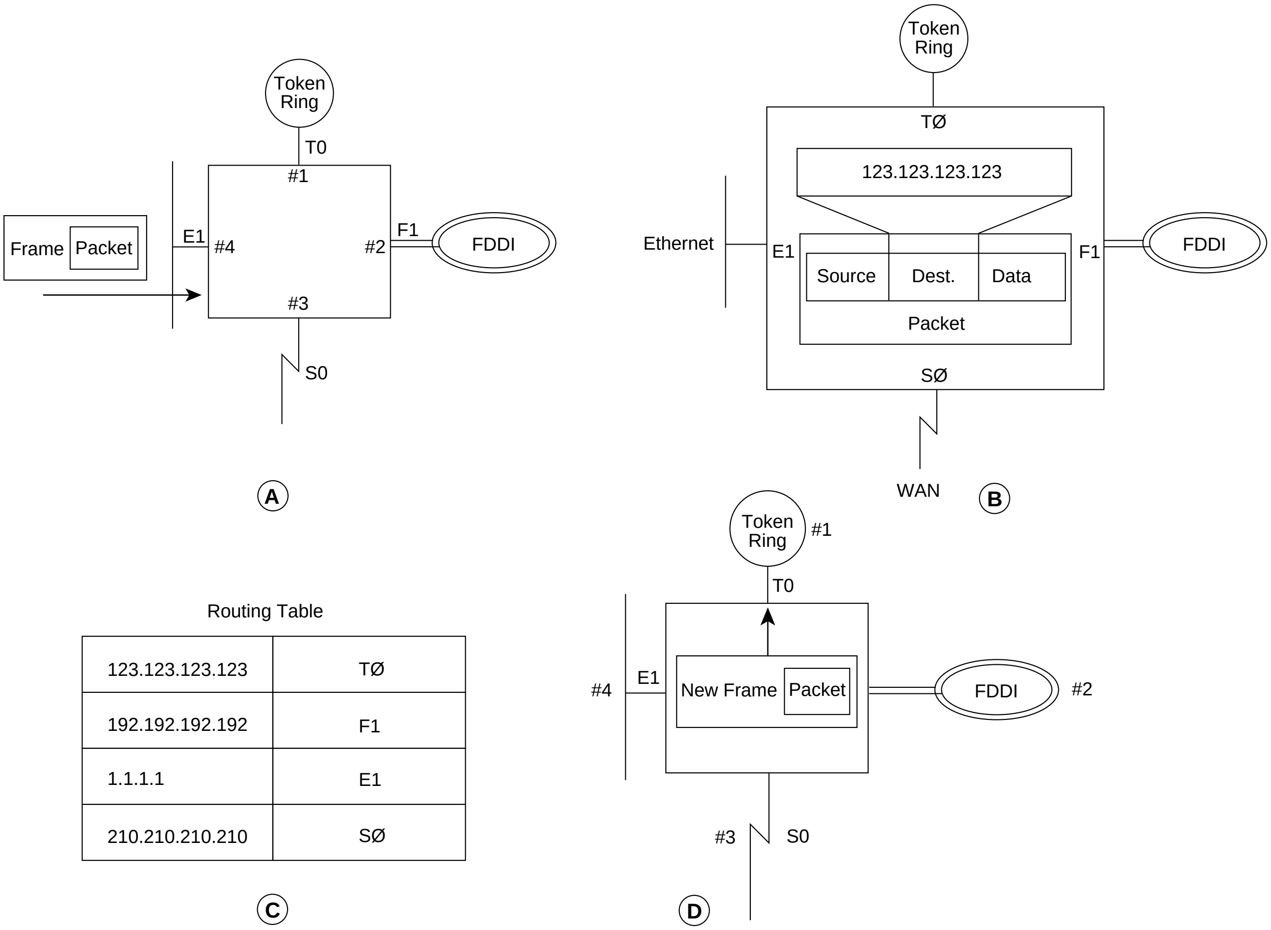


Switching Table

Frame Dest. Address	Output Interface
0001.2345.6789	2

Layer 3 switching looks at a network-layer, or protocol, destination address within a packet. The switch takes in the frame from the input interface, discards the layer 2 header portion, and examines the network-layer data in the packet. Layer 3 switching maintains a routing table and looks up the packet's protocol destination address. It then puts a new frame header back on the packet and sends it to the appropriate output interface. Figure 8 shows the layer 3 switching process, and Tables 6–8 show the routing tables for IP, IPX, and AppleTalk, and display how the packet goes out the appropriate output interface. Tables 1 and 2 show IP and AppleTalk ARP tables.

Figure 8. Layer 3 Switching Process



123.123.123.123	TØ
192.192.192.192	F1
1.1.1.1	E1
210.210.210.210	SØ

Table 6. IP Route Table

router# sho ip route	
Codes: C-connected, S-static, I-IGRP®, R-RIP, M-mobile, B - BGP, D-EIGRP, EX-EIGRP external, O-OSPF, IA-OSPF inter area, E1-OSPF external type 1, E2-OSPF external type 2, E-EGP, i-IS-IS, L1-IS-IS level-1, L2-IS-IS level-2, * - candidate default	
Gateway of last resort is 175.38.188.19 to network 200.170.51.0	
C	192.160.82.172 255.255.255.248 is directly connected, Ethernet2
D	192.160.82.56 255.255.255.248 [90/332800] via 192.160.82.186, 20:46:55, Ethernet6
S	192.160.82.232 255.255.255.248 [1/0] via 192.160.82.145
D EX	192.168.21.0 [170/293888] via 175.38.188.19, 00:10:34, Fddi0

Table 7. IPX Route Table

router# sho ipx route	
Codes: C-Connected primary network, c-Connected secondary network, R-RIP, E-EIGRP, S-static, W-IPXWAN connected, 16 Total IPX routes	
No parallel paths allowed Novell routing algorithm variant in use	
C	Net C0919A60 (SAP), is directly connected, 2480 uses, Ethernet7
C	Net C0919A61 (NOVELL-ETHER), is directly connected, 2480 uses, Ethernet7.1
C	Net AB685E11 (NOVELL-ETHER), is directly connected, 2480 uses, Ethernet1
R	Net C0919AF1 [1/1] via C0919A50.0000.0c0b.768b, 49 sec, 1 uses, Ethernet3
R	Net AB685EC0 [1/1] via C0919AB0.0000.0c37.1aa8, 16 sec, 1 uses, Ethernet0

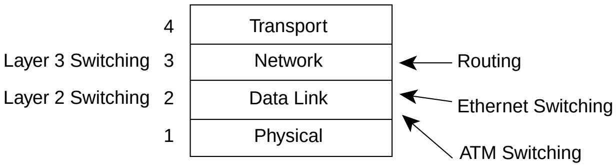
Table 8. AppleTalk Route Table

router# sho appletalk route	
Codes: R-RTMP derived, E-EIGRP derived, C-connected, A-AURP, S-static P-proxy, 864 routes in Internet	
The first zone listed for each entry is its default (primary) zone.	
R	Net 1-1 [4/G] via 6340.248, 4 sec, Fddi0, zone Christmastown
R	Net 4 [5/G] via 6340.248, 4 sec, Fddi0, zone Ethernet A5698327
R	Net 8 [2/G] via 6341.130, 2 sec, Ethernet0, zone Lab
C	Net 6359-6359 directly connected, Ethernet11, zone Lab

Switching and the OSI Model

Switching uses up to the first three layers of the OSI model, as shown in Figure 9. Ethernet switching operates at layer 2. It looks at a frame’s destination MAC address and moves it. ATM switching operates as an overlay technology at both layers 1 and 2. It looks at a cell’s destination address and sends it to the correct interface.

Figure 9. Switching and the OSI Model

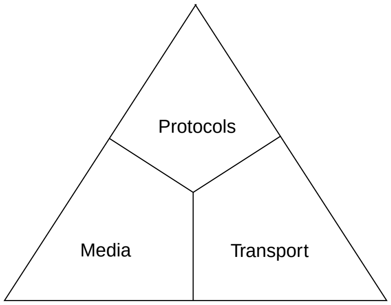


Routing operates at layer 3 of the OSI model. It examines packet information and forwards packets based on their network-layer destination addresses. Layer 2 switching can be related to Ethernet switching and ATM switching, and layer 3 switching supports router functionality.

Solving Network Problems

Switching and routing help to solve three main network problems: protocols, media, and transport. Protocol issues include the problems of broadcast scalability; media problems involve the collision domain; and transport problems include multimedia issues. We will discuss each problem in detail. Figure 10 displays these three problems.

Figure 10. Main Network Problems



Let's discuss how switching and routing resolve these three issues.

Protocols

Protocol problems include issues like broadcasts and broadcast scalability. A broadcast is a data packet that has a destination address of FFFFFF, which effectively means send it to everybody. Broadcasts are designed to be useful, but when many stations send out numerous broadcasts, the broadcasts can use up the bandwidth. This leaves less available bandwidth for stations to send and receive "real" data.

Routing can effectively handle these protocol problems at layer 3, while layer 2 switching cannot. When stations send out broadcasts, routing deals with the broadcasts at the network layer, not the data link layer. Routing controls broadcasts in several ways: with the use of routing tables, by way of proxy ARP default gateways or ICMP Router Discovery Protocol (IRDP), and network service tables using service proxies and service caching.

Routing Tables

First, routing caches the addresses of remote networks into routing tables. The next time a host broadcasts to obtain the address of a remote host, and the remote host's network address is in the routing table, the router responds on behalf of the remote host. The router then instructs the sending host to send future requests for hosts on that remote network to the router. This procedure is referred to as proxy ARP. If the router does not have the remote network in its routing table, it does not respond to the initial broadcast. If no routers have the remote network in their routing tables, then the packet is dropped. The router builds and maintains this routing table continuously. Proxy Arp applies only to IP. Table 9 shows an example of an IP Route Table.

Table 9. IP Route Table

router# sho ip route	
Codes: C-connected, S-static, I-IGRP, R-RIP, M-mobile, B-BGP, D-EIGRP, EX-EIGRP external, O-OSPF, IA-OSPF inter area, E1-OSPF external type 1, E2-OSPF external type 2, E-EGP, i-IS-IS, L1-IS-IS level-1, L2-IS-IS level-2, * - candidate default	
Gateway of last resort is 175.38.188.19 to network 200.170.51.0	
C	192.160.82.196 255.255.255.248 is directly connected, Ethernet7
C	192.160.82.180 255.255.255.248 is directly connected, Ethernet3
D	192.160.82.0 255.255.255.0 [90/40691200] via 192.160.82.178, 20:46:55, Ethernet0
S	192.160.82.232 255.255.255.248 [1/0] via 192.160.82.145

Table 9. IP Route Table

router# sho ip route	
D EX	198.92.27.0 [170/320000] via 175.38.188.19, 02:32:56, Fddi0

End stations can also be manually configured with their default gateway, so that the router address is known. This works until that router fails. Then the station loses connectivity to everything beyond that router. If the end station is configured for IRDP, however, the workstation knows when the router went down and finds an alternate router through hello messages that the routers send. In all three cases, the routers terminate broadcasts and reduce traffic on the network.

Network Services

In the second method for controlling broadcast radiation, routers on layer 3 devices cache advertised network services such as those used in IPX. When a router learns of a new network service, it caches the necessary information into its service table and does not forward broadcasts related to it. When a network service client sends a broadcast to locate that service, the router responds on behalf of that service and does not need to forward the broadcast. The router continuously updates these service tables, as shown in Table 10.

Table 10. IPX Service Table

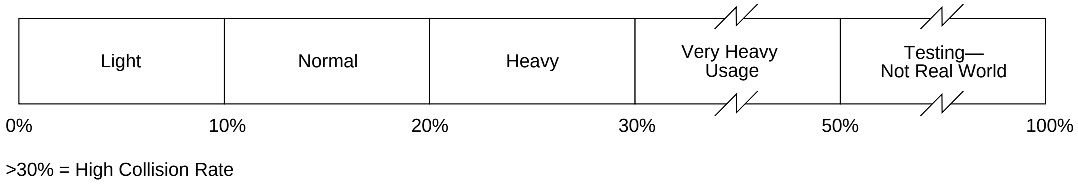
#sho ipx servers						
Type Itf	**Name**	**Net**	**Address**	**Port**	**Route**	**Hops**
P	4 ALASKA	871938.0000.0000.0001:0451		3/02	2	Fd1
P	4 WAN1	32.0000.0000.0001:0451		4/03	3	Fd1
P	4 LIBRARY	9999.0000.0000.0001:0451		4/03	3	Fd1
P	4 PRODUCTS	D.0000.0000.0001:0451		4/03	3	Fd0
P	4 DEMOROOM	3AE2DF85.0000.0000.0001:0451		5/04	4	Fd0

Layer 2 switching processes frames based on MAC addressing and propagates broadcasts; thus broadcast scalability does not improve with layer 2 switching.

Media

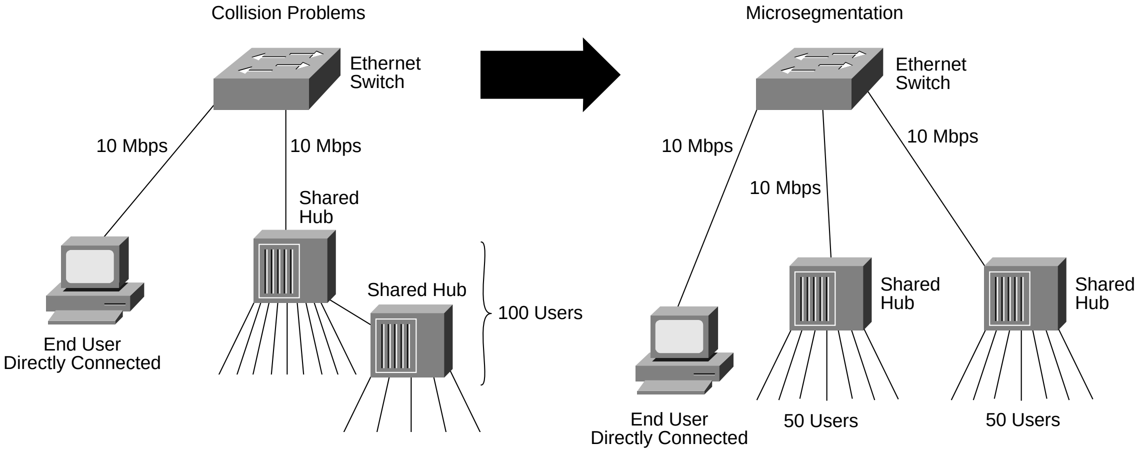
The media problem focuses on the collision domain, that is, collisions and collision avoidance. A collision domain can be described as a part of the network's bandwidth that users must share, also referred to as an Ethernet segment. As a device sends data out onto the segment, there is a chance that data from multiple users will "collide," because Ethernet does not guarantee delivery. When data collides, the data will have to be resent. The area where data can collide is called the collision domain. Ethernet switching improves collision problems that are commonly experienced in a shared-hub environment. Collision rates increase as segment utilization rises. In a network that has Ethernet connecting users to the wiring closets, a 30-percent utilization rate on a segment reflects heavy usage on that segment. Above 30-percent utilization, collisions start consuming more of the bandwidth than real data does. Figure 11 shows Ethernet utilization.

Figure 11. Ethernet Utilization



Utilization above 30 percent can occur when a segment has numerous users or when the stations on a segment run bandwidth-intensive applications. Ethernet switching, which is layer 2 switching, provides a solid solution for keeping collision rates low. Ethernet switches provide a dedicated 10/100 Mbps of bandwidth for each port. Each port on the switch represents a different segment. The collision domain is isolated on each segment. Therefore, for Ethernet switches, the collision domain is isolated to each port. Users that run bandwidth-intensive applications can be isolated on their own segments by connecting them directly to ports on the Ethernet switch. Or, as an increased number of users on a segment start to see a slower performance rate, they can be broken up and placed on two segments. This setup can be accomplished by using an additional port on the switch rather than buying additional devices such as bridges and concentrators, as is necessary with traditional networking equipment. Using microsegmentation to accommodate a growing network is illustrated in Figure 12.

Figure 12. Solving Increasing Network Size



By using another port on the switch, the collision domain is reduced, and each segment receives optimal utilization. Ethernet switches provide a dedicated 10 Mbps per port of bandwidth for user connections, and ofer both 10- and 100-Mbps trunk connections to connect to the network.

Transport

Transport problems, such as sending voice, video, and data over the network simultaneously, can be solved with ATM switching. Applications today are more bandwidth-intensive and require high transport speeds. Videoconferencing, imaging applications, and real-time interactive programs are some examples. Users are outgrowing shared-legacy LANs, in both utilization and speed.

ATM satisfies both the bandwidth and the speed requirements to make it the best solution for transport issues such as multimedia. ATM switching provides dedicated bandwidth of up to 155 Mbps today to each connection, guaranteeing a high-speed backbone. The dedicated connection eliminates the utilization and congestion concerns that come with shared bandwidth. The 155-Mbps speed is also fast enough for those applications that need more than a switched 10-Mbps Ethernet port at the desktop.

Campus Design Issues

This section covers areas to be considered when laying out a campus design and the priority of each. Topics to be covered include broadcasts, multimedia, VLANs, ATM/LAN Emulation (LANE), and backbone redundancy.

Broadcasts

Broadcasts in a network are essential traffic. Each protocol uses broadcasts to either discover routes or advertise services. IP clients use ARP to discover destination addresses, Novell servers use SAP to advertise its services, and AppleTalk runs Zone Information Protocol (ZIP) to send out information on the zones in the network. Also, network management applications use broadcasts for Simple Network Management Protocol (SNMP) queries of devices running on the network.

Two broadcast terms are particularly important: broadcast storms and broadcast radiation. Broadcast storms refer to a situation in which the number of broadcasts on the network severely degrades the network's performance. Broadcast radiation refers to the broadcast/multicast traffic on a network segment that an end station must process, even though the host may not benefit from it.

Ideally, broadcast storms are not common and do not last long. Because they bring the network down, the misbehaving device hopefully is discovered quickly and removed from the network. It is easier to find the broadcasting device in a hierarchical network than in a flat network. A hierarchical network with routers limits the broadcast domain, allowing initial problem isolation automatically. In a flat network, the broadcasts propagate across switches and bridges, making the broadcast domain much larger. Isolating the problem device becomes much more difficult in this situation.

Broadcast radiation is not monitored closely in a network, nor are its effects on end stations easy to monitor. It affects the end stations' CPU performance and takes longer to detect than broadcast storms. Factors such as the operating system, network interface cards (NICs), and CPU processors all play a part in how a station handles broadcasts. For example, running the IP protocol on a SunSPARC2 and a SunSPARC5 produce very different results. Table 11 compares Sun and PC Pentium performance numbers.

Table 11. CPU Utilization of End Stations in Relation to Broadcast Radiation

SunSPARC2 Running SunOS 4.1.3		SunSPARC5 Running Solaris 2.4		PC Pentium 120 MHz with 3Com Fast Etherlink PCI Ethernet Adapter (Windows 95 Kernel Processor Usage Application)	
Broadcasts/Second	**CPU Utilization**	**Broadcasts/Second**	**CPU Utilization**	**Broadcasts/Second**	**Processor Utilization**
100	3%	100	3%	100	2%
1000	23%	1000	10%	1300	9%
3000	69%	3000	28%	3000	25%

To monitor and evaluate how broadcast radiation affects every workstation is not an easy task. Broadcast radiation is not protocol-dependent either. For example, stations running IPX or IP will process IPX, IP, and AppleTalk broadcasts, whereas stations running AppleTalk process only AppleTalk broadcasts.

It is useful to remember that, although switches will improve the collision domain compared to shared hubs, they will not improve the broadcast domain. If stations begin to experience performance problems, but the collisions remain low, broadcast radiation may be one cause of the problem.

Multimedia Design Issues

Another significant design consideration is multimedia. Point-to-point multimedia is not really an issue in the overall scheme of the campus design, except for bandwidth requirements. With point-to-point multimedia, the necessary bandwidth must be available to both points. Point-to-multipoint multimedia, however, is a major design consideration. Point-to-multipoint multimedia is a broadcast frame at layer 2. For IP, the destination layer 3 address is a Class D IP multicast frame with a multicast address.

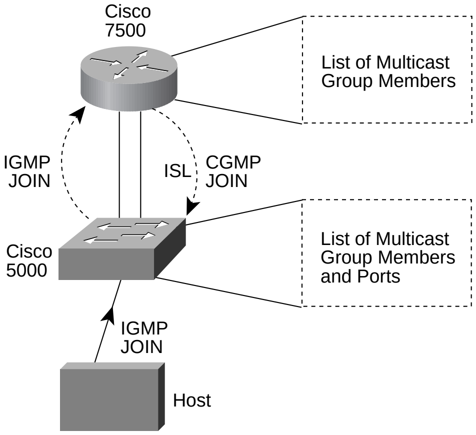
For large-scale, or even small-scale operations, this kind of framing can be a major issue in switched networks. Today, all stations in the broadcast domain (all layer 2 stations on connected switches for a flat topology or every VLAN member for VLAN organized switches) receive this multimedia point-to-multipoint broadcast. Current Motion Picture Experts Group 2 (MPEG2) transmission has a rate potential of 6 Mbps. If we are designing a two-video transmission network with MPEG2, then when we transmit the second video, all other activity on the network ceases on the Ethernet switches, because they are busy handling the 12 Mbps of video. Users who are not using the video have no bandwidth to run their applications.

Typical multimedia applications today do not run at the 6-Mbps rate, but 1.5 Mbps is not uncommon. Distance learning, videoconferencing, and video-on-demand applications operate at 1.5 Mbps or more. With two video streams, for example CNN at work and a videoconferencing application, more than 3 Mbps of bandwidth is easily used up at each switch port and end station. Using 30 percent of an Ethernet segment just for two broadcast streams takes up a lot of the usable Ethernet bandwidth. For users who are not interested in those multimedia applications, the applications use up valuable bandwidth, and their end-station application performance degrades with no benefit to them.

Cisco is actively looking to solve this issue and is making an effort to separate point-to-multipoint multimedia from standard switching fabric. When point-to-multipoint multimedia is required today, either the Catalyst™ 1200 or Catalyst 2000 switches can offer solutions. The Catalyst 1200 has some IP routing capabilities. It supports Internet Group Management Protocol (IGMP); therefore, it can filter the multicasts for ports that do not initiate an IGMP JOIN. The Catalyst 2000 switch supports multicast registration. It has the capability to register MAC-layer addresses; it also lists the ports to which these packets are to be forwarded and disables the normal flooding of unregistered multicast packets on a per-port basis. This capability increases network performance for users who are not using the multimedia applications.

A software solution that Cisco is working on is the Cisco Group Management Protocol (CGMP) for the Catalyst 5000 and router platforms. The router creates a CGMP JOIN message when it receives an IGMP JOIN message from a host. The CGMP JOIN message includes the MAC addresses of the multicast group. The router then multicasts the CGMP JOIN to 0100.0cdd.dddd. Each switch that receives the CGMP JOIN message adds the port on which the host resides to the bridge-forwarding table entry for the group. With CGMP, the switch forwards the multimedia multicast to only the ports that belong to that multicast group. Figure 13 shows an example of CGMP.

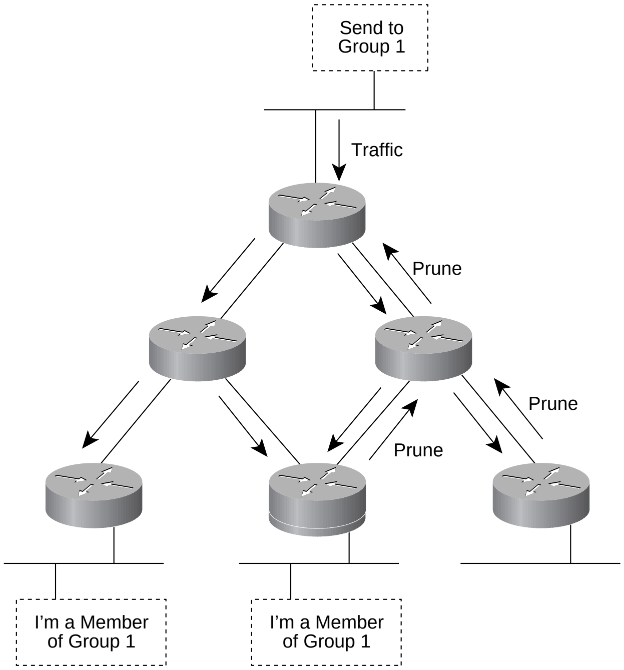
Figure 13. CGMP



Cisco also has software solutions at the router level available now: Protocol Independent Multicast (PIM) offers two types of multipoint traffic distribution patterns to address multicast routing scalability; they are dense mode and sparse mode.

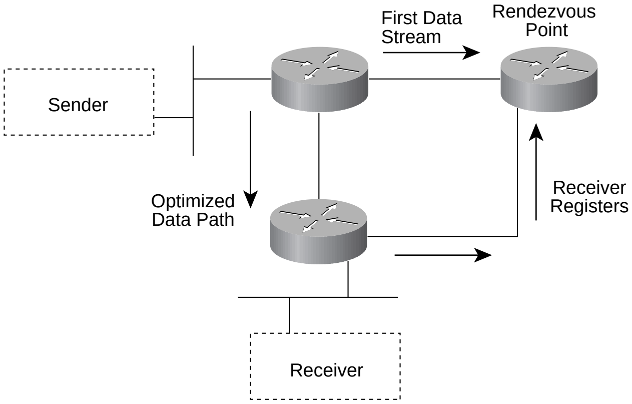
Dense-mode PIM is most appropriate in a campus design situation when there are few senders and many receivers, and the multicast traffic is constant and high volume. Dense-mode PIM utilizes Reverse Path Forwarding; it will flood the network and prune back where there are no multicast group members. In networks where there are members on most subnets or VLANs, dense-mode PIM is effective. Figure 14 shows dense-mode PIM.

Figure 14. Dense-Mode PIM



In networks where there are only a few multicast receivers, or the traffic is intermittent, sparse-mode PIM may be a better solution. Sparse-mode PIM utilizes a rendezvous point. Multicast senders transmit to the rendezvous point first, and receivers register with the rendezvous point first. Once the data begins to go from sender->rendezvous point -> receiver, the routers optimize the data path for the shortest route. At this point, the traffic does not need to go through the rendezvous point at all. Figure 15 displays sparse-mode PIM.

Figure 15. Sparse-Mode PIM



For multimedia traffic in the network, the routers can utilize PIM between the routers and CGMP between the switches and routers.

Multimedia requirements are also discussed in the "VLAN Design" section of this paper. More detailed information about multimedia networks can be found in the Networked Multimedia Design Guide located on the Web at Universal Resource Locator (URL) wwwin.cisco.com/mkt/data/product/entepris/atech/index.htm.

Campus Mobility

Network clients are becoming increasingly mobile. Users are dialing in from homes, airports, customer sites, and so on. AppleTalk and IPX easily support mobile users by their nature. The IP protocol is more difficult; two methods, VLANs and DHCP, support mobile users. DHCP is more specialized for mobility, whereas VLANs solve additional problems as well as mobility.

VLAN Design

To VLAN or not to VLAN, that is the question. In this section, we will look at the advantages and disadvantages of VLAN architecture in a campus LAN design.

VLAN Advantages

VLAN design offers easier moves and changes in a network design than traditional networks. You can create VLANs for separate functions, and through software configurations you can easily move users to different LAN segments. It is easy to physically move a user; you need only to configure the new switch port to the correct VLAN through network management and connect the user after the move. Other advantages of VLANs include:

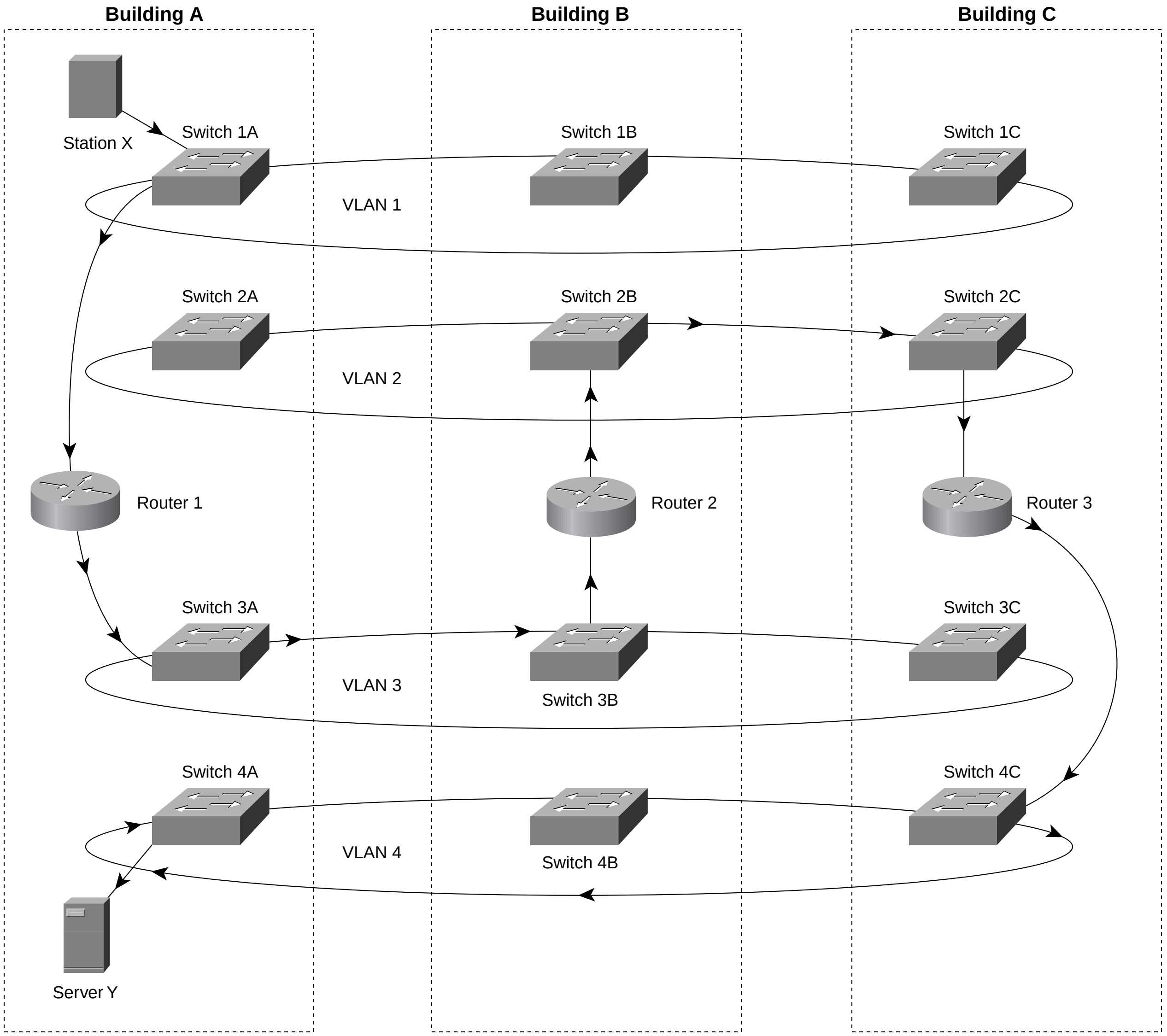
- Broadcast Control—VLANs offer smaller broadcast domains than flat networks, which can substantially reduce broadcasting on local segments. VLANs do not communicate with other VLANs except through routing; therefore broadcasts do not propagate across VLAN domains.
- Physical Location Independence—VLANs dispel the physical constraints of traditional networks. Users on a subnet no longer need to be in the same physical vicinity; they can spread across the campus.
- Security—VLANs allow high-security projects to be isolated. If there is no router connecting a high-security VLAN to the rest of the network, there is no way to communicate with that VLAN to get the classified information.
- Performance—VLANs can also separate bandwidth-intensive projects to be placed on a separate VLAN, so that other users will still have high performance at their stations.

VLAN Disadvantages

Although VLANs have some definite advantages, they also have disadvantages, such as:

- Troubleshooting—The first disadvantage of VLANs is more difficult debugging; an example is shown in the topology in Figure 16.

Figure 16. Possible VLAN Data Paths



It is beneficial to have multiple routers connected to a single VLAN. When this design is used, layer 2 switching takes precedence over layer 3 frames in most cases, so IP ARP packets are answered by the first router to respond, not necessarily the closest router. Macintoshes in the network use the first RTMP router they hear from. What this architecture does is unlink the physical topology from the logical topology.

Let’s look at a “worst-case” scenario. This design emphasizes the importance of accurate campus network design when using VLANs.

Looking at Figure 16, if Station X on VLAN 1 in Building A wants to access data from server Y on VLAN 4 in Building A, the following data path would occur:

1 Station X sends request out to switch 1A in Building A.

2 Switch 1A forwards it to router 1.

3 Router 1 routes it from VLAN 1 to VLAN 3 and sends it to switch 3A, still in Building A.

4 Switch 3A sends the request to switch 3B in Building B.

5 Switch 3B forwards it to router 2 in Building B.

6 Router 2 routes the request from VLAN 3 to VLAN 2 and sends it to switch 2B in Building B.

7 Switch 2B sends the request to switch 2C in Building C.

8 Switch 2C forwards it to router 3.

9 Router 3 routes the request from VLAN 2 to VLAN 4 and sends it to switch 4C in Building C.

10 Switch 4C forwards the request to switch 4A in Building A.

11 Switch 4A directs the request to server Y.

When server Y responds, it takes the same data path to get back to station X.

Adding more router physical connections to the VLANs may not mean a shorter logical data path. If the same router paths respond first, the data path does not change. This scenario is acceptable when the system works without failure, but if a router or end station sends corrupt information, debugging this problem is much more difficult. When first introduced, someone commented that VLANs were like AppleTalk zones; this analogy might be truer than we had expected.

- Redundancy—The second disadvantage of VLANs is that high-speed redundant links are hot backups, not usable networks. Layer 2 topologies must use spanning trees to discover loop-free environments. When redundancy is created, layer 2 devices cannot use them for data transmission, only for hot backup of the primary spanning tree. Layer 3 devices (routers) can use redundant media for data transmission.
- Traffic Patterns—A third disadvantage of VLANs relates to the traffic pattern. VLANs work well in a well-behaved network; that is, if 80 percent of the traffic is local in the VLAN (intra-VLAN), and only 20 percent needs to go to another VLAN (inter-VLAN). With this in mind, centralized server network designs cannot easily implement VLANs without redesigning their networks. The VLAN architecture must have both local servers and centralized servers in the network for optimum performance.

DHCP

VLAN technology is not the only technology that can perform logical networking and support moves and changes in the network. A possible alternative to VLANs is Dynamic Host Configuration Protocol (DHCP), a protocol that allows clients to connect to networks and have IP addresses dynamically assigned to them. DHCP supports adds, moves, and changes in the network. This protocol may solve your needs in an IP environment, and you may not need VLAN technology.

DHCP provides configuration parameters to Internet hosts. DHCP consists of two components: a protocol for delivering host-specific configuration parameters from a DHCP server to a host, and a mechanism for allocating network addresses to hosts.

DHCP is built on a client/server model, where designated DHCP server hosts allocate network addresses and deliver configuration parameters to dynamically configured hosts. DHCP supports three mechanisms for IP address allocation: dynamic allocation, automatic allocation, and manual allocation.

With automatic allocation, DHCP assigns a permanent IP address to a host. With dynamic allocation, DHCP assigns an IP address to a host for a limited period of time. Using manual allocation, the network administrator assigns the host’s IP address, and DHCP is used to convey the assigned address to the host. Dynamic allocation supports mobility effectively, therefore we will introduce it now.

Dynamic Allocation

Dynamic allocation is a mechanism that provides easy adds, moves, and changes in the network. It can be run with or without VLANs. Dynamic allocation allows automatic reuse of an IP address. As a client disconnects from the network, the IP address goes back into the pool of available addresses for the server to reassign to another client.

With DHCP, users can now move locations and connect back into the network easily. DHCP works well with users who both telecommute and work in the office. They may be moving their stations two or three times a day. DHCP dynamic allocation easily adds them into the network, regardless of where they are physically located.

DHCP dynamic allocation works well for stations that stay temporarily connected to the network. The DHCP servers keep a pool of IP addresses that are shared among a group of temporary hosts. DHCP servers have permanent IP addresses; only the clients have dynamic addresses.

DHCP is designed for IP; IPX and AppleTalk automatically use a dynamic address configuration, so in a network that uses more than one protocol, DHCP will solve the IP configurations.

Most modern IP stacks include support for a DHCP client, such as *Cisco TGV,* Microsoft, FTP, and NetManage. DHCP servers are available for Windows NT and UNIX.

VLANs have the disadvantage of being very difficult to debug, but DHCP takes a large administrative effort to set up and maintain. The administrator must set up both the servers and the routers to support DHCP. Thus it is a trade-off for the network administrator whether to use VLANs or DHCP. Refer to RFC 1541 for more information on DHCP.

VLAN Strategies

For users who want the advantages of VLANs and can live with the downside, and for whom DHCP will not answer all their needs, we will investigate the issues of VLAN design.

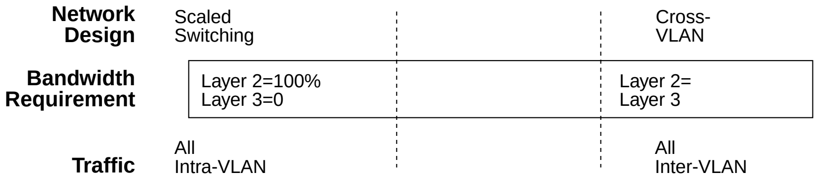
Let’s discuss a few VLAN strategies, such as:

- Per-protocol VLANs
- Per-department VLANs
- Micro-VLANs (many very small VLANs)
- One to five VLANs for a large campus

Each of these strategies has an upside and a downside. Let’s start with an overview of VLAN strategies. Your VLAN strategy is inherently linked to your specific campus network and the data flow, server placement, and quantity of cross-VLAN traffic. By definition, cross-VLAN traffic and enterprise traffic need to be switched at layer 3, probably by a router Therefore, any VLAN strategy that has very a small number of users in each VLAN, or where traffic is primarily crossing VLANs, means that this campus must have a healthy router content. Again, the ideal VLAN network should have 80-percent intra-VLAN traffic and 20-percent inter-VLAN traffic. The more inter-VLAN traffic there is, the more routers are needed to process the data traffic.

In the extreme, where all data is cross-VLAN or enterprise, the routing bandwidth (layer 3 bandwidth) must equal the switching bandwidth (layer 2 bandwidth). Although these networks are rare, they can be easily designed by the novice LAN administrator. An example of this design is the per-department VLAN strategy where one of the “departments” is the centralized data center where all the company’s servers exist. Figure 17 shows bandwidth requirements of layers 2 and 3 in relation to network traffic flow.

Figure 17. Layers 2 and 3 Bandwidth Requirements



Even if you understand the network flow today, it can change dramatically overnight with new applications and technologies. The impact of Web servers on most of our networks is a good indicator that nothing is stagnant. Layer 3 bandwidth can be economized by reducing or eliminating inter-VLAN traffic and enterprise-wide traffic. Most network administrators have little or no control over the traffic patterns of their networks.

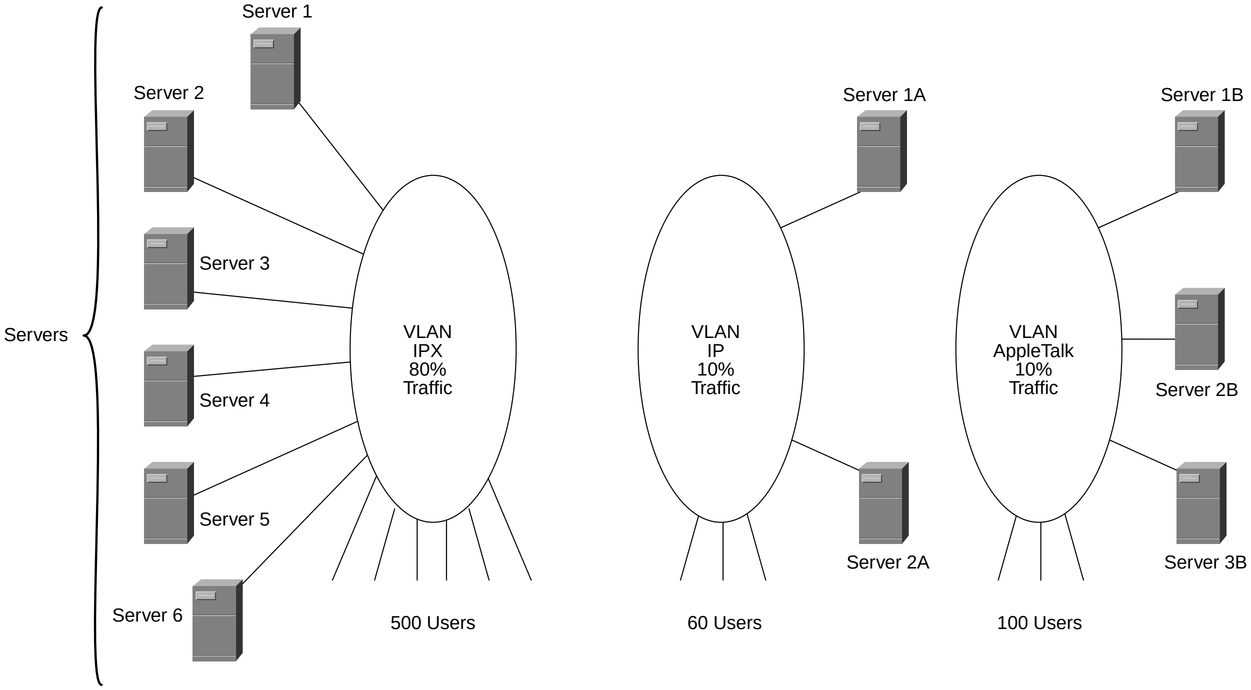
The key issue of VLAN design is adequate routing content to support the design of your VLANs. If you have adequate routing content, your strategies will work. If you underpower your topology with inadequate routing capability, the network will show signs of stress. A VLAN under stress is paradoxically demonstrated by router stress at the VLAN connection point. Underpowered VLANs offer 100-percent media load to the router where drops caused by burst conditions occur. These devices may demonstrate 100-percent CPU loading because of broadcast constraints, because all broadcasts must be assessed by each router, and large-scale VLANs with more than 200 users can generate healthy broadcast storms.

Let’s review some VLAN strategies now.

Per-Protocol VLANs

Designing a VLAN network based on protocols may work in smaller networks that run nonroutable protocols. For large networks that use one protocol much more than others, this design does not scale well. Imagine trying to put IPX on one VLAN, with IP and AppleTalk on two others. If IPX is 80 percent of the traffic on the network, the users will most likely be impacted by the server broadcasts, which occur every 60 seconds. There are no routers to cache the advertised services and reduce the amount of broadcasts. Nor can the routers limit the broadcast domain when the protocol is all on one VLAN. This VLAN design may work with small networks that have balanced traffic between protocols. Figure 18 shows a logical representation of a VLAN network separated by protocols.

Figure 18. Protocol VLANs



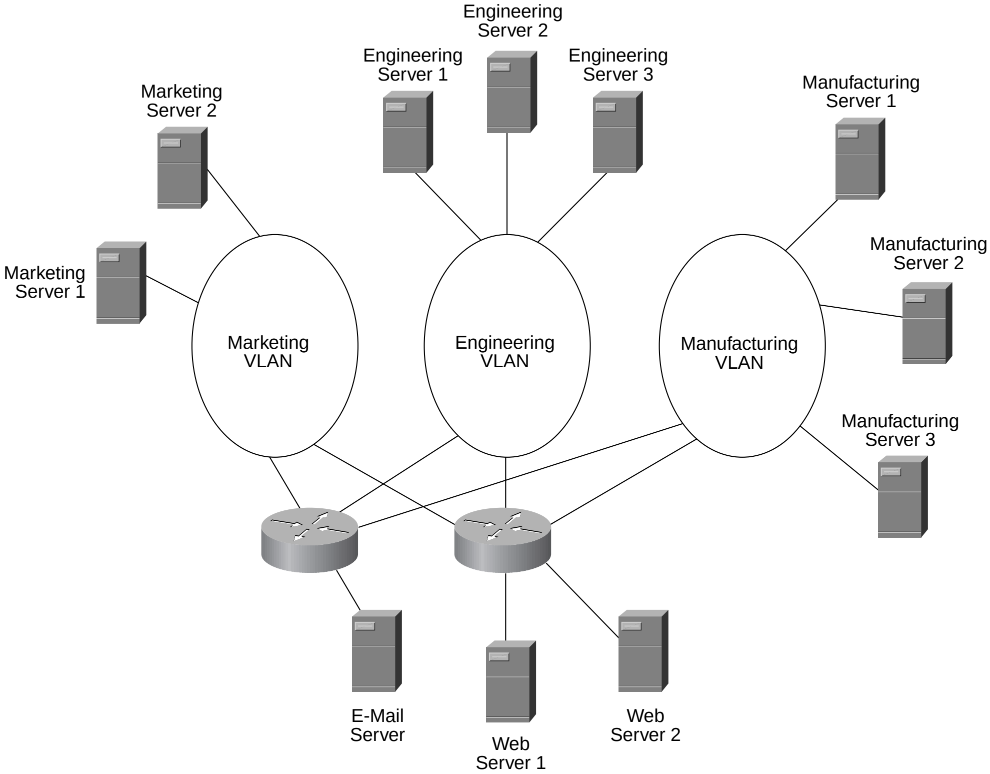
This design also assumes that workstations are running only one protocol. For stations running more than one protocol, this VLAN strategy will not work. The end station would have to have intelligence about what VLAN it is on, rather than the NIC or the switch port having this intelligence. This technology will not be available in the near future. NICs also do not have VLAN intelligence now (except for ATM LANE). Today, the station would have to have a separate NIC for each protocol it is running, and each card could be on a different VLAN because it would connect to a different switch port. The switch port would allocate the VLAN ID that the NIC card belongs to.

Per-Department VLANs

Designing VLANs on a per-department strategy is perhaps the most common method of subnetting a network. There are engineering servers, marketing servers, and so forth. Without multiple VLAN support on the NIC, however, the traffic pattern surely does not fit the 80/20 ideal. To access an engineering server from the marketing VLAN creates inter-VLAN traffic. The location of the servers in this strategy also has an impact on inter-VLAN traffic. If the servers are all centralized, then more routers are needed to support the traffic that is going inter-VLAN from each department to the servers. The best solution for departmental VLANs is to have local department servers located on each VLAN, and central servers to support things **(traffic? software?)** such as e-mail, Meeting Maker software, and centralized Web servers.

Currently, ATM with LANE can support multiple VLANs on a single NIC. The central servers could participate in several VLANs in this situation. If the central servers can only support one VLAN, the core section of the network must be carefully designed to avoid major bottlenecks and supply enough bandwidth for the inter-VLAN traffic. Figure 19 shows an example of departmental VLANs.

Figure 19. Departmental VLANs

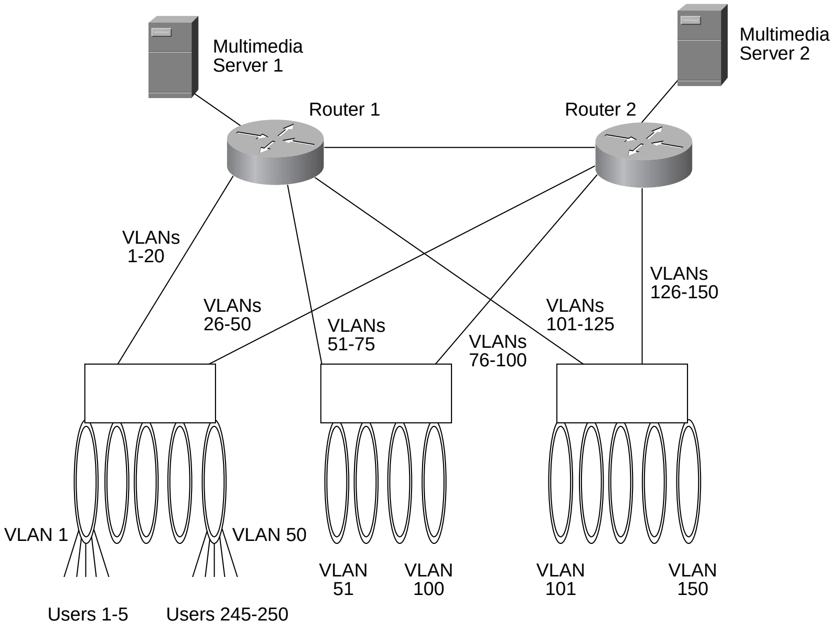


Micro-VLANs

Micro-VLANs (VLANs that have only a few members) support videoconferencing applications well. There will be a large number of micro-VLANs in a network. For networks that run multimedia applications on a regular basis, this may be the best VLAN design today. Video applications send multicasts, and switches handle multicasts the same as broadcasts; the switch sends the multicast out all its ports. By using micro-VLANs, the bandwidth of video applications (which can be as high as 6 Mbps) do not affect a large number of users who are not using the application. If a VLAN is receiving the multimedia multicast packet, each station that is a member of that VLAN must process the multicast packet. With larger VLANs, this situation can create broadcast radiation problems, as discussed earlier in this paper.

This micro-VLAN design certainly has limitations when regarding data transmission, however. Because of the micro-VLAN design, there must be a large amount of router bandwidth to support the inter-VLAN traffic, and because of the small number of members in each VLAN, most traffic is inter-VLAN. Instead of the traffic being 80/20, it looks more like 20/80, with 20 percent of the traffic being intra-VLAN, and 80 percent inter-VLAN. It is even worse if there is no local server on the micro-VLAN; then there is even more inter-VLAN traffic. Figure 20 shows an example of micro-VLANs.

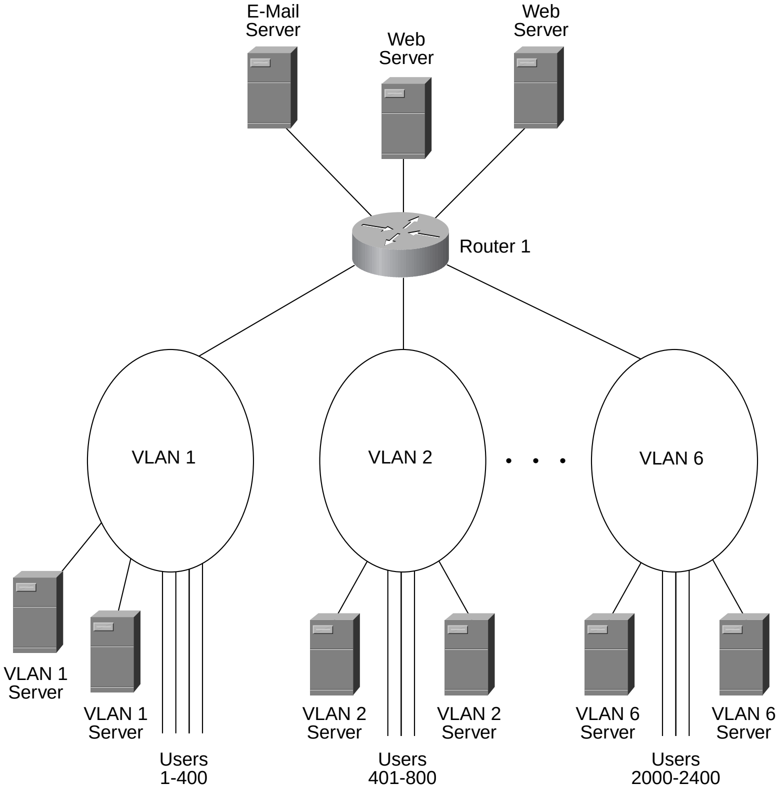
Figure 20. Micro-VLANs Example



Large VLANs

The final VLAN strategy we will consider is large VLANs, VLANs that have numerous members. If a campus is designed on this strategy, there will be a small number of VLANs, each of which contains numerous members. This VLAN strategy should fit into the 80/20 traffic pattern easily. To incorporate this VLAN strategy, the network traffic pattern must be well known. This traffic pattern will aid in determining how many users can be on a single VLAN. Figure 21 shows a large VLAN design solution.

Figure 21. Large VLAN Example



Broadcast traffic also plays a part in this VLAN strategy. Because there are numerous members in each VLAN, broadcast traffic may become significant. Each member will see and process all the broadcasts because there are no routers within the VLAN to suppress them.

The large VLAN strategy also assumes that servers are local to each VLAN. The inter-VLAN traffic should consist only of clients accessing a "central" serve, one that communicates with several or all the VLANs. It is not important to know which VLAN a central server belongs to; it is more important to know how much traffic it is handling.

The large VLAN strategy works when the traffic is almost all data. For high multimedia application traffic, this strategy is not recommended; the micro-VLAN strategy may be more appropriate.

Network Considerations

When deciding on a VLAN strategy, potential users should investigate the following network characteristics:

- What is your network traffic pattern? Can you separate it into local and centralized traffic?
- How large a broadcast domain can your network/VLAN tolerate?
- How many protocols are running, and what are the characteristics of each?
- Is your network critical enough that it *must* be running 24 hours a day?
- What kind of applications are most used, and are they broadcast-multicast-intensive or mostly unicast?
- How is your current network subnetted, and what problems is this causing?

VLAN Issues

Cisco supports three VLAN architectures; ATM LANE, 802.10 over Fiber Distributed Data Interface (FDDI), and Inter-Switch Link (ISL) over Fast Ethernet. Each architecture has issues to consider. A discussion of the issues of each architecture follows.

ATM Design Issues

Let's look at using ATM in a campus LAN today. The three forms of ATM networking are Native Mode, Classical IP, and LANE; on the horizon is Multiprotocol over ATM (MPOA). We believe that MPOA as a standard is 12 to 18 months away. Since Native Mode ATM does not support current layer 3 protocols as they exist today, we can eliminate it from a discussion of enterprise networking. Classical IP is limited to IP only and does not fit most of our customers' profiles. So in general, 90 to 95 percent of our customers who use ATM in their campus networks will be using LANE. Much investment is being made in ATM technology, and there has been progress. LANE is the only VLAN standard today that has achieved interoperability between vendors, and the added bandwidth to 155 Mb is an advantage. However, LANE nullifies all the ATM advantages such as quality of service (QOS). ATM also is not cost-effective for wide deployment, such as to the desktop.

For more detailed information on ATM, refer to the Designing Switched LANs document, located on the WEB at wwwin.cisco.com/mkt/data/product/enterpris/atech/index.htm, or the "Internetworking with ATM" white paper by Anthony Alles.

802.10 over FDDI

Cisco developed VLANs over FDDI by using 802.10. Although 802.10 is a security standard, not a VLAN standard, Cisco has implemented the VLAN strategy into it. In time, VLANs over FDDI will migrate to 802.1q as that becomes an industry standard. FDDI is a mature, stable technology, but it is bandwidth-limited. FDDI switching is not viable because the LightStream® 2020 does not support 802.10.

ISL over Fast Ethernet

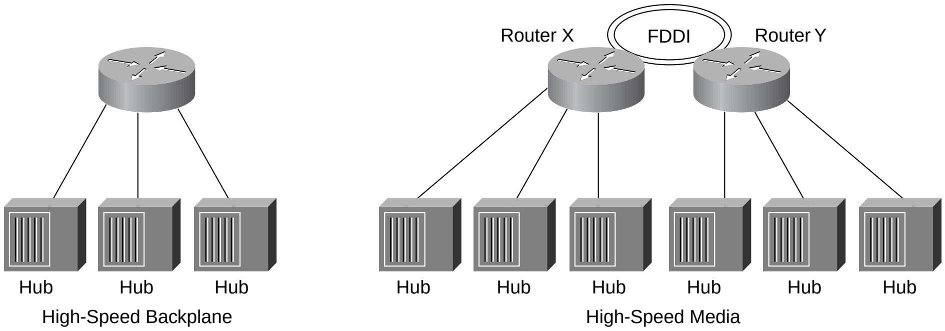
Cisco supports VLANs over Fast Ethernet by utilizing its proprietary ISL. Fast Ethernet provides higher bandwidth and is cost-effective. ISL will be easy to modify in order to support Gigabit Ethernet in the future. However, because ISL is proprietary, VLAN interoperability is limited to only Cisco switches and routers. Cisco is working with companies to broaden the support of ISL, such as the agreement with Intel.

Backbone Redundancy

Redundancy in a network is an important part of a network design. Some networks need to function for part of the day, and others are working 24 hours each day. Still other networks are so critical that they need to be fully functioning all the time. Each backbone needs to support redundancy in different ways.

A network backbone traditionally has consisted of one of two items: high-speed backplane or a high-speed media, as displayed in Figure 22.

Figure 22. Network Backbone



Network redundancy started with Ethernet. Users could connect to an Ethernet hub, and the hubs connected together, all on the same Ethernet segment. When networks needed more than one shared Ethernet, routers took over the backbone. Multiple Ethernet networks could connect to the router, and the router’s backplane processed all the data.

As Ethernet became too slow, the backbone swung back to a high-speed media with FDDI. Multiple routers and concentrators could connect to the 100-Mbps backbone. When ATM was introduced as the new high-speed technology, a single ATM switch became the backbone. The ATM switch fabric handles each 155-Mbps connection internally. As ATM networks grow larger, several ATM switches and ATM routers will comprise the network’s backbone.

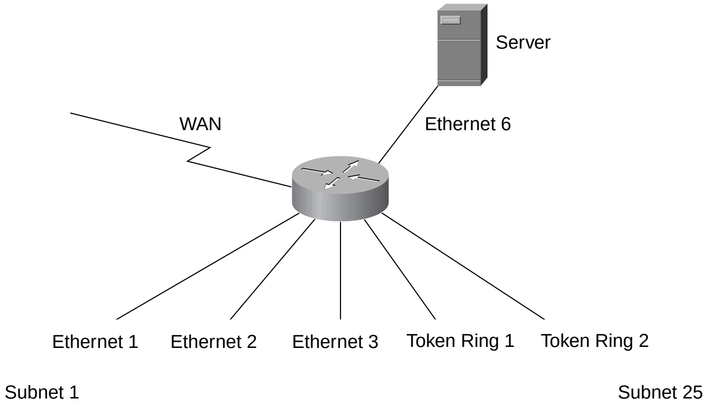
Another option for fully redundant high-speed backbones is available: the core mesh matrix, a matrix that takes the network out of the traditional high-speed backplane/media loop.

Each of these backbones is discussed here in terms of its redundancy level; the high-speed backplane, high-speed media, and the core mesh matrix.

High-Speed Backplane

In a traditional network design where the network backbone is actually one device such as a router, the backbone consists of the high-speed backplane. Figure 23 shows a typical high-speed backplane backbone.

Figure 23. High-Speed Backplane



In this scenario, the power and hardware of the device need to have redundancy. Within the backbone device, redundant power supplies, fans, and possibly interface cards can be put in to guard against specific hardware failures. If the backplane of the primary device fails, however, not all the internetwork traffic is routed. Traffic is limited to the local subnet, and any traffic trying to reach the central server or the WAN would fail as well.

With ATM becoming more popular today, the ATM switch can make up the high-speed backbone. The backplane has much more speed but no more redundancy benefits than a router.

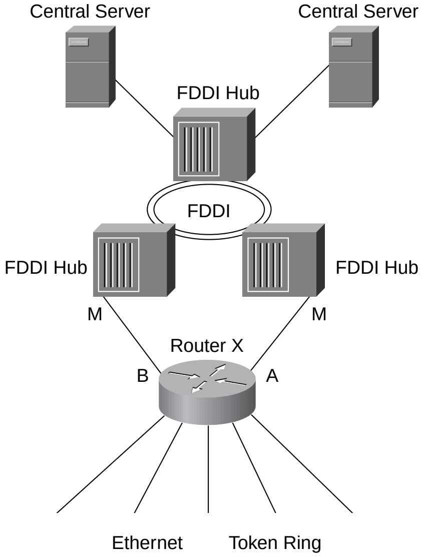
There can be a redundant ATM switch in case of a switch failure. However, with ATM and its connection-oriented traffic, incorporating redundant interfaces is more difficult to accomplish than traditional connectionless technologies. Each device that requires redundancy in ATM needs an additional interface, whether it is a router, workstation, or LAN switch with an ATM uplink. With the cost of ATM today, high redundancy may not be worth the price.

LAN switches are easier because they support spanning tree. Having redundant links to the same device is easy; two connections to the same device use only an additional switch port. When an interface fails, spanning tree automatically moves the blocked port into forwarding mode.

High-Speed Media

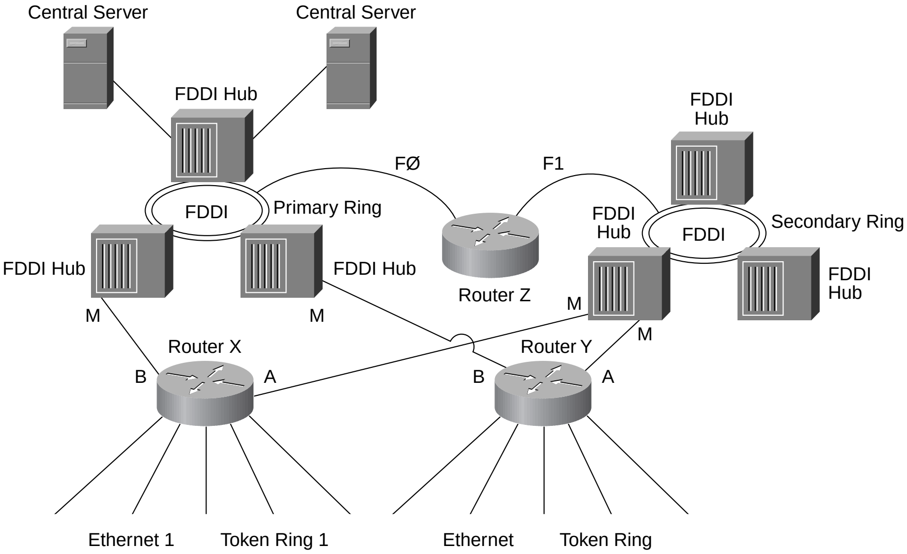
In a different traditional network design, the backbone consists of high-speed media such as FDDI. In addition to hardware redundancy, there must be redundant data paths into and out of the backbone. Figure 24 shows an FDDI high-speed media backbone.

Figure 24. High-Speed Media Backbone



An FDDI ring with dual attachment station (DAS) connections is inherently redundant and fault-tolerant, offering an automatic backup ring through dual homing. There can also be a backup Ethernet backbone; it is relatively easy to install the backup FDDI path. Stations connected to the devices on the FDDI backbone can be dual-homed, making one path the primary path and the other path secondary, as illustrated in Figure 25.

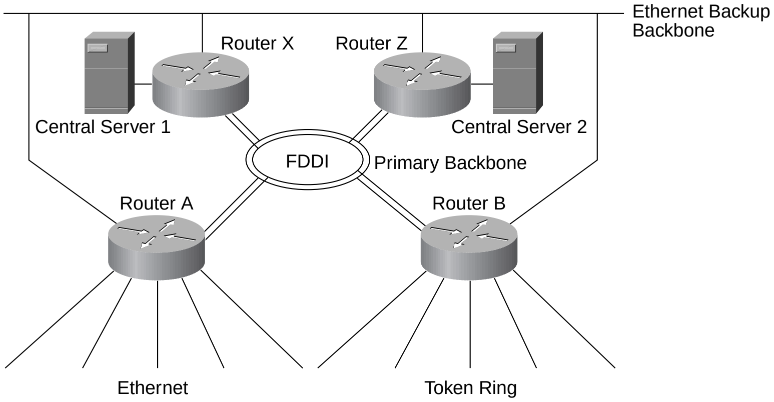
Figure 25. Data Path Redundancy



The data path is protected by redundancy from the moment it gets onto FDDI to the time it is routed back to Ethernet (or Token Ring).

Another common design is to have a backup Ethernet backbone. As networks have migrated over to FDDI, existing Ethernet backbones have been kept in place as standbys in case the FDDI backbones failed. This solution was inexpensive and redundant. Figure 26 gives a design example of an Ethernet backbone backup.

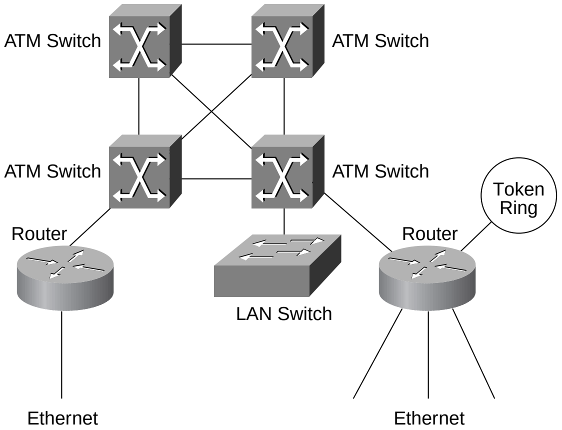
Figure 26. Ethernet Backbone Backup



The largest problem with this design today is that the network is dependent upon 100 Mbps in the backbone. If the FDDI backbone fails and the Ethernet backbone comes up, the network gets a tenfold hit. Data moves from a 100-Mbps token orientation to a 10-Mbps collision orientation. Not many networks can handle that much of bandwidth hit.

In today’s designs, ATM backbones are becoming increasingly popular. With an ATM backbone consisting of more than one switch, a level of redundancy can be incorporated. There can be multiple paths across ATM switches to reach the same endpoint, as shown in Figure 27.

Figure 27. ATM Backbone



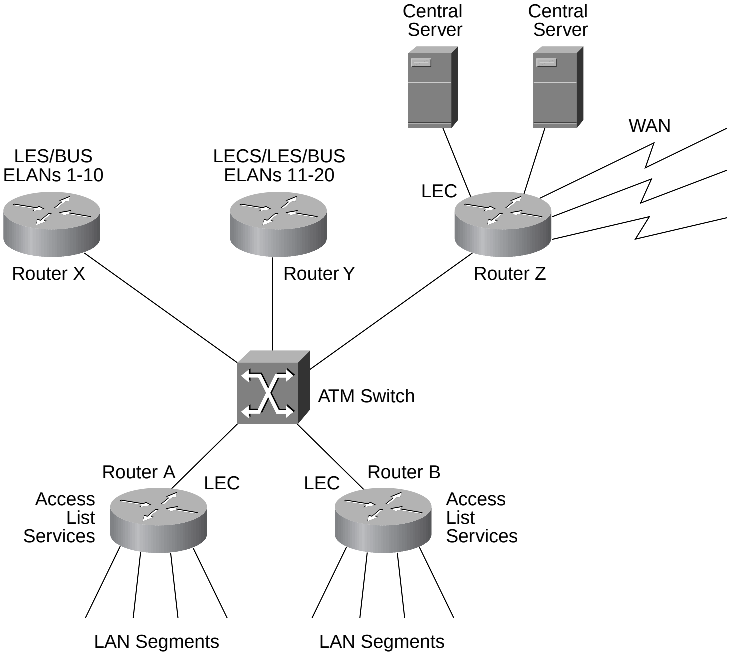
Even though ATM is connection-oriented, if a connection between switches is lost, ATM can reestablish another connection by going through a different switch. There are, however, some single points of failure with ATM and LANE.

The first is the LAN Emulation Configuration Server/LAN Emulation Server/broadcast and unknown server (LECS/LES/BUS) architecture. There is only one LECS operating as the master at any given time. Several LECs can be configured, so that if the master LECS fails, the next LECS address can activate. This solution is a Cisco proprietary redundancy solution; therefore it will work only with Cisco ATM LANE devices. Other vendors' ATM equipment cannot support multiple LECSs, so if their LECS fails, no new ATM connections can be established.

Next is the LES/BUS, a standard design that supports one LES/BUS pair per emulated LAN (ELAN). Two possible points of failure may occur here: First, if the LES/BUS fails in software, the ELAN that the LES/BUS supports stops working; and second, multiple LES/BUS pairs can be configured on one router or Catalyst 5000 ATM interface. If that ATM interface fails in hardware, then all the LES/BUS pairs stop working as well. In this case, multiple ELANs are brought down. Cisco is currently working on LES/BUS redundancy; however, this solution again will be proprietary. Other vendors do not have LES/BUS redundancy.

When designing the LANE network, care must be taken to prevent configuring too much on one interface. The LES/BUS services for all the ELANs should be separated between several devices, depending on how many ELANs are created. With routers, other traffic that the router is handling besides LECS/LES/BUS services must be considered, and whether other devices can offload that traffic. (See Figure 28.)

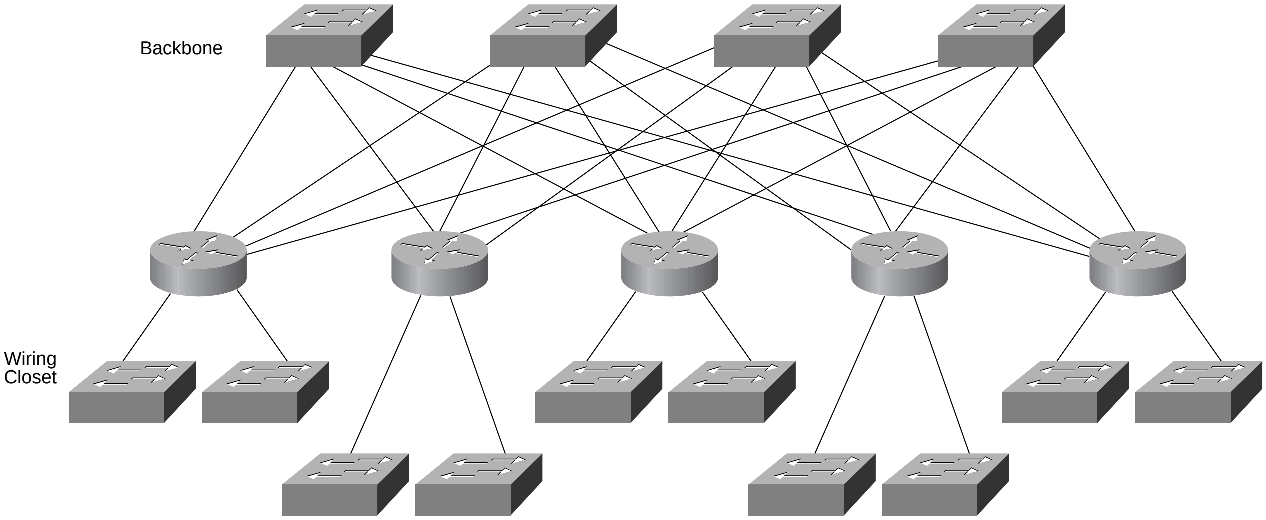
Figure 28. Router Functions



Core Mesh Matrix

Some networks cannot experience any failures at all. Full-blown redundancy and protection are among the highest priorities when designing this kind of network. To have a fully redundant backbone so that there is no single point of failure, the backbone must be complex and consist of a mesh of paths to the core backbone. Figure 29 displays a campus core router mesh.

Figure 29. Campus Core Router Mesh



In this scenario, there is no single point of failure. There is a redundant path as well as redundant hardware for every device in the backbone. The core consists of high-speed media such as ATM or Fast Ethernet and the central core is made up of high-speed switches. At the outer core, routers are performing layer 3 switching functions and determining the best path to the destination. The edge devices are layer 2 switches that connect the users to the redundant backbone.

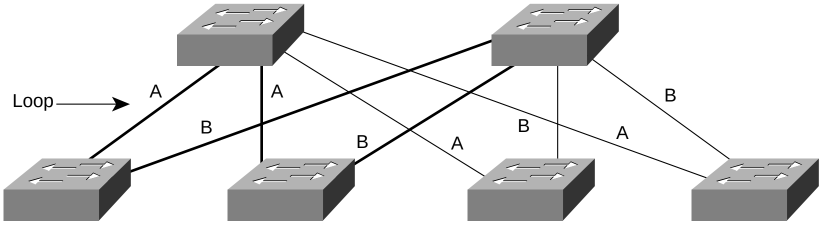
This design supports full redundancy for networks that absolutely cannot have any downtime. It is the most complex and most expensive of the redundancy solutions, but provides the greatest fault tolerance.

Data Path Redundancy

Redundancy can also be viewed in terms of traffic direction. Layer 2 switching uses hot backup with spanning tree for redundancy, and Layer 3 switching uses parallel paths.

With hot backup, the backup data path is blocked until the primary path goes down. (See Figure 30.)

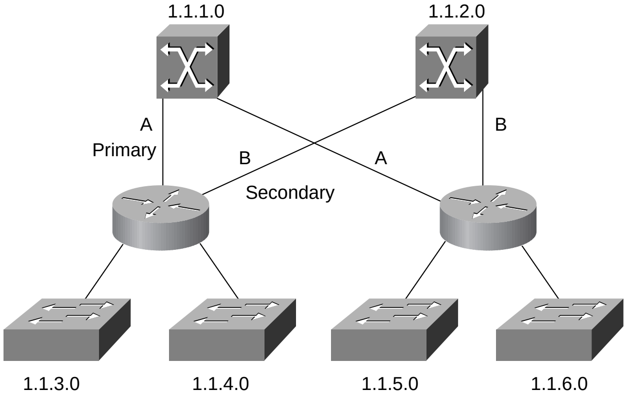
Figure 30. Hot Backup Redundancy



Spanning tree removes the loops in the network topology, so if data path A is operating, then data path B will be in blocking mode. If data path A fails, then data path B goes from blocking mode to forwarding mode. There is one spanning tree, which operates at layer 2, per VLAN domain.

With parallel paths, both paths are active, and the layer 3 mechanism determines the best path to take. (See Figure 31.)

Figure 31. Parallel Paths



With parallel paths, there is 2 x 155 (or 100) Mb of effective, usable bandwidth. Hot backup allows only 1 x 155 (or 100) usable bandwidth. However, layer 3 is more complex, and can cause slower network performance. The router determines the best path to take to reach the destination. If the best path fails or is congested, then the other path is used.

Designing Campus LANs

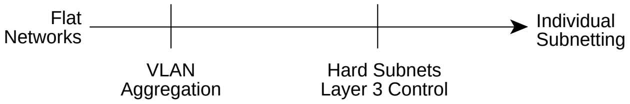
The way you design your network depends upon your considerations of the previous discussions about campus design issues: broadcast problems, traffic utilization, network problems, multimedia, VLANs, ATM/LANE, and redundancy issues. We describe one possible solution for each size network in the sections that follow. These solutions are not the only solutions, but in each case the solution is one of many. The correct solution for any network depends on the importance of the factors mentioned previously to the customer's network goals.

There are several ways to look at designing a campus LAN. In the examples that follow, we approach the design based on the number of users on the network, which also relates to the amount of traffic on the network. We divide campus LANs into four categories:

- Small Networks: fewer than 100 users
- Medium Networks: 100 to 500 users
- Large Networks: 500 to 2000 users
- Enterprise Networks: >2000 users

First let's look at general network designs. Networks can range from completely flat to thousands of individual subnets, or fall somewhere in between. Both extremes are unreasonable, and most networks fall in the middle somewhere. The two middle scenarios we will discuss are VLAN aggregation and hard subnets utilizing layer 3 control. VLAN aggregation is closer to the flat network theory, and hard subnets are closer to the complete individual subnet theory, as shown in Figure 32.

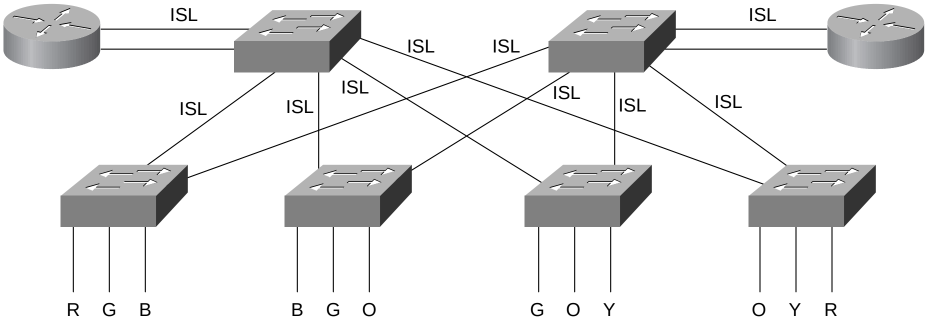
Figure 32. Network Design Range



VLAN Aggregation

VLANs operate optimally with large, flat subnets, using routers only to pass inter-VLAN traffic. Figure 33 shows a general physical topology of a network using VLAN aggregation.

Figure 33. VLAN Aggregation

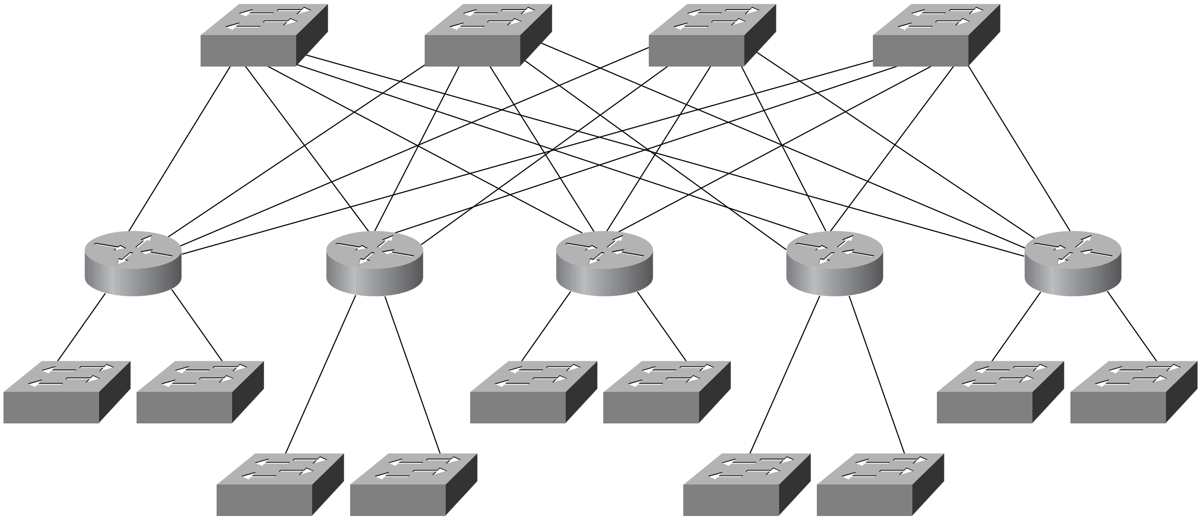


Most of the traffic travels across switches. Traffic does not pass through a router at all unless it requires some layer 3 function, so all the data that is intra-VLAN passes through high-speed switches only. The throughput is fast because there are no complex computations to perform at layer 3. The routers can have multiple physical connections to the switches to provide more bandwidth or to load-balance the inter-VLAN traffic. Each connection can be configured for every VLAN, with different priorities set to help balance the traffic.

Hard Subnets

Hard subnets require more layer 3 activity, a better solution if the network is running DHCP instead of VLANs. Figure 34 shows a physical topology of a layer 3 controlled network.

Figure 34. Enterprise Campus Network



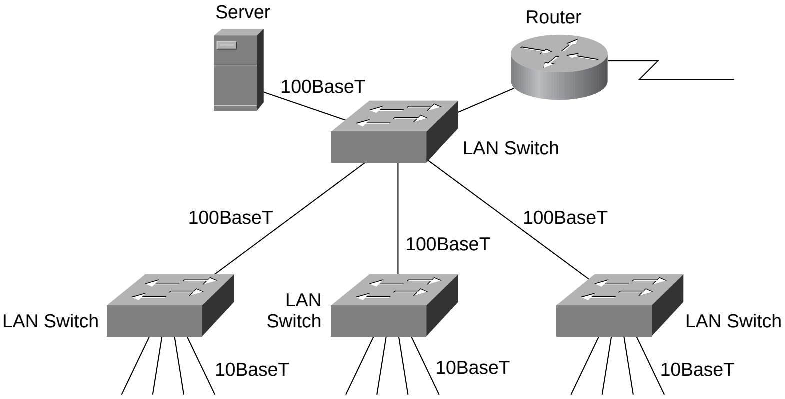
By using more layer 3 control, the network can use more security through the routers. There is more packet manipulation in this topology, but it is necessary if security is an issue. Most traffic passes through routers, where the security for the network is configured. The routers also pass the traffic between different subnets. If multimedia applications are widely used, this network design is preferred, because the multicasts and broadcasts will impact fewer users than in the VLAN aggregation design.

Now let’s look at campus designs based on the size of the network. We will begin with the small network, and move up to the enterprise network.

Small Networks

For small networks, the campus design can be simple and straightforward. With less than 100 users, the network can consist of one LAN (VLAN) with minimal subnetting. The network consists mostly of switches, with possibly a router to perform CGMP functions or to connect to the Internet. This design is commonly referred to as scaled switching, as shown in Figure 35.

Figure 35. Small Campus Network

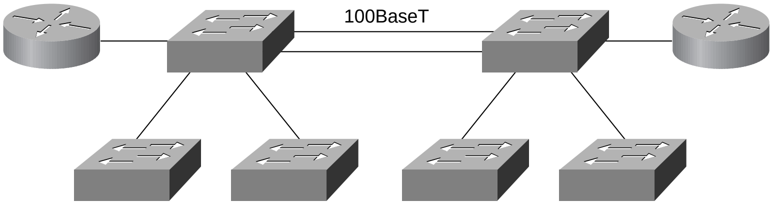


The network can be microsegmented with switches so that bandwidth does not become a limitation. The few servers on the network can be connected to high-speed interfaces on the switch. There should not be a problem with broadcasts or traffic congestion, and multiple VLANs are not needed for a network of this size. Redundancy is probably not high on the priority list and may consist of having some spare devices to turn on in case of a hardware failure. If multimedia is a factor in this network, utilizing CGMP on the router and switches or subnetting the network may be a solution.

Medium Networks

As the small network increases in size to 100 to 500 users, the system administrator should consider subnetting the network. The same network architecture can be used for the medium network as for the small network. Users connect to switch ports by 10 Mb, or 100 Mb for high-bandwidth machines, and these switches connect to the core switch via a high-speed link. The router connects to the core switch as well, and provides connections to the Internet. The router performs all the data transfer between different subnets. If the network spans more than one building, the same physical architecture can be duplicated in the other locations. The core switches can be connected together, forming the high-speed backbone. As more traffic flows through the router, more routers can be added to increase performance or to add more security to the network. Figure 36 shows a two-building, medium-sized campus network.

Figure 36. Medium-Sized Campus Network



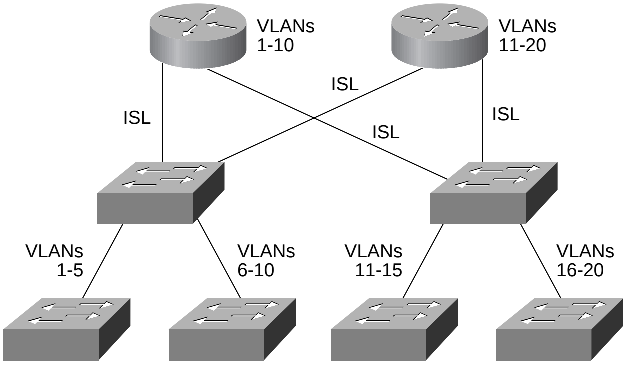
Central servers can connect to the backbone switches, and local servers can connect to the workgroup switches. In this network, VLANs can be introduced, and DHCP is also successful. By subnetting the network, multimedia multicasts can work successfully if the router and switches implement CGMP. If CGMP is not implemented, then the multicasts will propagate throughout every subnet if there is even one station listening to it. This scenario could create severe broadcast propagation, which could be reduced by utilizing CGMP and PIM.

Large Networks

When networks reach 500 to 2000 users, they need to spread geographically, and they must be more carefully planned out. Instead of adding switches and routers as the need arises, the whole network must be considered when expanding. Subnets, traffic patterns, and broadcast domains must be tracked accurately in order to predict network behavior. The speed and fault tolerance of the backbone become more important factors in the design as well. We will discuss two network designs: one focuses on VLAN implementation, and the other one concentrates on DHCP.

By implementing VLANs on a network that has grown from medium to large size, it is important to know the traffic patterns of the network. If the traffic can fit into the 80/20 rule of VLANs, 80 percent local to the VLAN and 20 percent inter-VLAN, then implementing VLANs will be successful. Figure 37 shows a network design for VLANs.

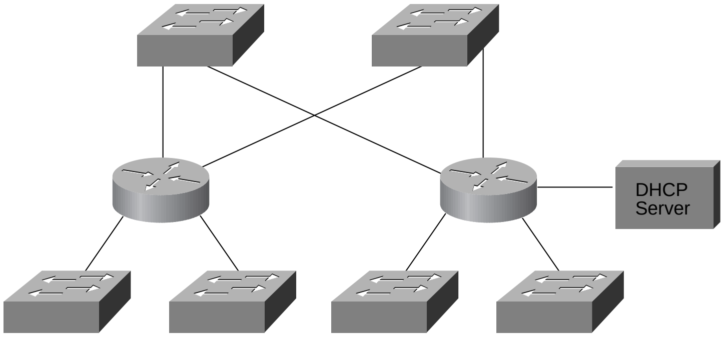
Figure 37. Large Campus Network—VLANs in the Workgroup



Note how the routers are located one level away from the workgroup. For VLANs, the goal is to have all the traffic traverse the switches via a high-speed link, and only inter-VLAN traffic to travel up to the router. Ideally, this traffic will be low, less than 20 percent, and will not suffer a performance hit from the router. Central servers, which need to send traffic to multiple VLANs, will be located either on the switches or on the router. Inter-VLAN traffic needs to go through the router anyway, so the location of the central server is flexible. Local servers should be connected to the switch via a high-speed connection. If the local server can be connected to a workgroup switch, then that reduces the amount of traffic going up to the core switches.

The second network design we will discuss is the network that utilizes DHCP. Figure 38 shows an example of this network.

Figure 38. Large Campus Network—DHCP for the Enterprise



The switches and routers are reversed in this network design. The switches are one level higher than the routers. The DHCP server is connected to a router. As users connect to the network, the traffic goes through the router to the DHCP server, and the server assigns an address to the new host. This network is appropriate when most of the traffic crosses subnets. VLANs will not work well, and DHCP is more flexible in this scenario. In this design, the traffic crosses the router and goes into the backbone consisting of switches. With more routers, the traffic is more easily handled, and there is less chance of network congestion. Local servers are still located at the workgroup switch, but central servers should be connected to the backbone switches.

Both of these networks are easy to migrate to from the medium-sized network design. As the network grows, more equipment is needed to handle the traffic. The basic network infrastructure should not have to be completely changed, however.

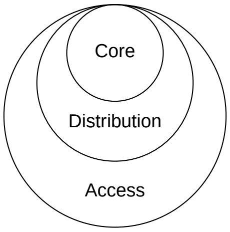
Enterprise Networks

When a network becomes greater than 2000 users, it can become very complex and demanding. It must be designed and maintained continuously, and the whole network must be studied when changes or expansions need to occur. There should be a well-formed hierarchy within the network, and redundancy is most likely an important consideration. We will discuss the enterprise campus network here. This solution is not the only one but we will show you how it is easy to migrate from a large network to an enterprise network without having to restructure the network. Let's discuss the layers of hierarchy in a network first.

Access/Distribution/Core Hierarchy

In a well-formed hierarchical network, there should be three easily defined layers, traditionally referred to as the access, distribution, and core layers. (See Figure 39.)

Figure 39. Access/Distribution/Core Hierarchy



Each of these layers provides a different function. The layers do not need to exist in clear and distinct physical entities, but the functionality needs to exist in an enterprise network. To aid in the understanding of these functional layers, we have modified the traditional layers to access/workgroup, distribution/policy, and core/backbone.

The main function of the access/workgroup layer is to connect users. Other functions represented by this layer are shared bandwidth, switched bandwidth, MAC-layer filtering, and microsegmentation. LAN switches are most commonly seen in this layer of the network.

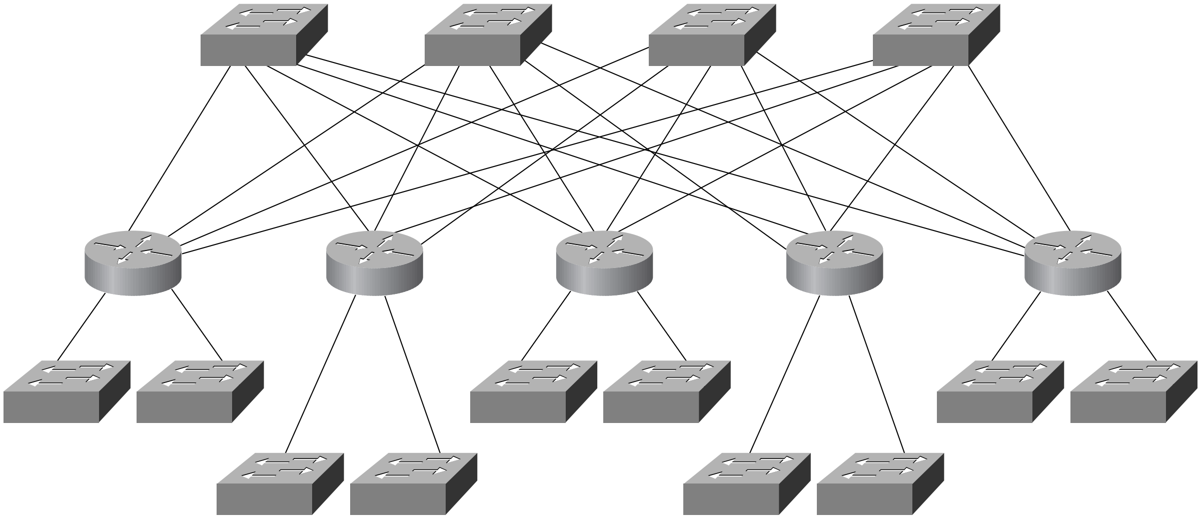
The distribution/policy layer performs the policy-based operations. Thus it performs the complex, CPU-intensive calculations such as filtering, access lists, inter-VLAN routing, GGMP, broadcast/multicast domain definition, and address or area aggregation. This layer may also contain the local servers. Routers reside in the distribution layer, and LAN switches may reside here as well.

The core/backbone layer is the backbone of the network. It should be high-speed and concerned mainly with switching traffic as quickly as possible. It should not get involved in "expensive" packet manipulation. ATM, Fast Ethernet, or FDDI should make up the core backbone. The central servers may also be attached to the high-speed backbone in the core. ATM switches, high-speed routers, and LAN switches can be found in the core.

Enterprise Network Design

The enterprise network design shown earlier in Figure 34 supports hierarchical levels. LAN switches are at the access/workgroup level, consistent with the smaller network sizes previously discussed. Routers are in the distribution/policy layer, and switches are at the core/backbone. Figure 40 shows an enterprise network design.

Figure 40. Enterprise Campus Network



This network design also has a high level of redundancy. There is no single point of failure in this design until we reach the desktop connection. If a workgroup switch fails, those users connected to it are down, but no one else in the network suffers.

Also, Figure 40 broken down into sections is just several of the large network designs meshed together. Instead of having only a few switches and routers, there are many more of them, spanning several buildings. The WAN connections can also come into the routers, and the data is processed the same as the local traffic. The central servers are connected to the core switches via high-speed connections, and the local servers are connected to one of the workgroup switches.

If the network is utilizing VLANs instead of DHCP, then the routers and switches may be swapped, to have a routing core for inter-VLAN routing, and the switches handle most of the traffic. If the enterprise network is running multimedia applications, then it is vital that the routers run PIM, CGMP, and other QOS functions such as Resource Reservation Protocol (RSVP) or custom queuing.

Summary

We have discussed a variety of campus network design issues in this paper, and we have made it clear that network administrators need to consider and plan for many more factors than in the past. They need to baseline not only their backbone traffic patterns, but also the bandwidth that user end stations need.

Because of the new applications such as multimedia that users are demanding today, broadcast domains must be carefully planned and managed to maintain high-performance networks. Decisions about campus mobility and whether the answer is DHCP or VLANs must also be made. Backbone technology (Fast Ethernet, ATM, FDDI) and the level of backbone redundancy the network needs have also become complex issues.

We have discussed all these areas and have offered some possible solutions. These solutions are certainly not the only ones available; network designers must decide on solutions that best meet their specific network needs. This paper also shows that there are no single solutions to satisfy network requirements. By looking at the complete enterprise campus network instead of individual network sections, network designers will be able to accurately define their current network problems, as well as plan for future growth and user requirements.

CISCO SYSTEMS

Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
World Wide Web URL:
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

Cisco Systems has more than 125 sales offices worldwide. To contact your local account representative, call Cisco's corporate headquarters (California, USA) at 408 526-4000 or, in North America, call 800 553-NETS (6387).
0896R

0896R