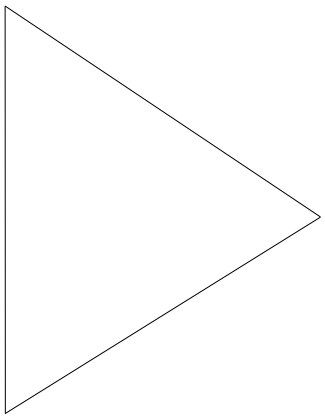
Design Implementation Guide



Designing Switched LANs

Introduction

New client/server applications have driven the need for greater bandwidth in traditional shared-media environments, and LAN switching is being deployed to address this requirement. The current generation of LAN switches are inherently layer 2 devices, and, although they are replacing and enhancing the shared-media concentrators and hubs in use today, they are not replacements for routers. LAN switches as layer 2 devices cannot alleviate the necessity of providing layer 3 routing functionality.

In a flat bridged network all broadcast packets generated by any node in the network are sent to and received by all other nodes in the network. The ambient level of broadcasts generated by the higher layer protocols in the network—known as "broadcast radiation"—will typically restrict the total number of nodes that the network can support. In an extreme case the effects of broadcast radiation can be so severe that an end station spends all of its CPU power on processing broadcasts.

Virtual LANs (VLANs) have been designed to address two problems: the scalability issues of a flat network topology and the simplification of network management by facilitating network reconfigurations (moves and changes). A VLAN consists of a single broadcast domain and solves the scalability problems of a large "flat" network by breaking a single bridged domain into several smaller bridged domains or VLANs.

At a higher level many customers see the promised ease of configuration of VLANs as a means of reducing the perceived complexity of maintaining their current routed networks. This perception arises from an incomplete understanding of the function and value of routing in their networks. Routers provide critical services, such as broadcast filtering, security, address summarization, and traffic flow management, as well as intra-VLAN routing. All these services are useful, and necessary in building a scalable, stable network.

Several key issues need to be considered when designing and building switched LAN internetworks. This paper will look at VLANs, at the role of routing in a campus environment, and at general network designs that successfully implement switched LANs and VLANs.

The paper is organized into sections that discuss the following topics:

- The benefits of VLANs
- The role of routing in a switched network
- Network designs in switched networks and scalability issues
- Various technologies used to implement VLANs, including 802.10, Inter-Switch Link (ISL), and LAN Emulation (LANE)
- The effects of broadcast traffic in a switched internetwork
- An overview of Cisco's network management for administering VLANs
- An overview of Cisco's various switch architectures and their relative positioning

Why Implement VLANs?

The incentives for customers to use VLANs are appealing; the promised ease of configuration (dynamic moves and changes) coupled with higher bandwidth for individual users (switched LANs) is very attractive. VLANs not only offer these benefits, but they are crucial for building scalable switched internetworks.

A VLAN consists of many end systems, either hosts or network equipment (such as bridges and routers), connected by a single bridging domain. This bridging domain is supported on various pieces of network equipment (for example, LAN switches) that operate bridging protocols between them, with a separate bridge group for each VLAN. First-generation VLANs are based upon various layer 2 bridging and multiplexing mechanisms—for example, IEEE 802.10, LANE, and ISL—that allow the formation of multiple, disjointed, overlaid broadcast groups on a single network infrastructure.

VLANs are used to connect a set of related users, regardless of their physical connectivity. They could be located across a campus environment or even across geographically dispersed locations. The users might be related in terms of all being in the same department or functional team, or the data flow between them might be such that it makes sense to group them together. The power of VLANs comes from the fact that moves and changes can be achieved simply by configuring a port into the appropriate VLAN. Expensive, time-consuming recabling to extend connectivity in a switched LAN environment is no longer necessary, because network management can be used to logically “drag and drop” a user from one VLAN to another.

VLAN Features

- *Broadcast Control*

  Just as switches isolate collision domains for attached hosts and forward only appropriate traffic out a particular port, VLANs refine this concept further and provide complete isolation between VLANs. A VLAN is a bridging domain, and all broadcast/multicast traffic is contained within it.

- *Security*

  VLANs provide security in two ways. With the first method, high-security users can be grouped into a VLAN, possibly on the same physical segment, and no users outside of that VLAN can communicate with them. In addition, because VLANs are logical groups that behave like physically separate entities, interVLAN communication is achieved via a router or a switch (bridge). If interVLAN communication occurs via a router, all the traditional security and filtering functionality that routers provide can be used, because routers are able to look at Open System Interconnection (OSI) layer 3 information. If bridging is being used for interVLAN communication (as in the case of nonroutable protocols), only the filtering capabilities of traditional bridges can be used (OSI layer 2). In the second scenario, where bridging occurs between VLANs, in essence one big VLAN is created.

- *Scalability and Performance*

  The logical grouping of users allows, for example, an engineer making extensive use of a networked CAD/CAM station or testing a multicast application to be put into a VLAN containing just the engineer and the servers he needs. His work does not affect the rest of the engineering group. This setup improves performance for the engineer by putting him on a dedicated “LAN” and helps performance for the rest of the engineers by not affecting their work.

- *Network Management*

  The logical grouping of users, divorced from their physical or geographic locations, allows easier network management. Pulling cables to move a user from one network to another is no longer necessary; moves and changes can be achieved via software.

The Role of Routing in Switched Networks

Ongoing zealotry abounds in the debate on the relative merits of routing or bridging (now referred to as switching). We believe that both of these technologies address a particular set of problems and that regarding them as complementary is essential when designing a switched network. We shall examine the merits of each and consider the solution space that each technology addresses.

A close examination shows that the nature of LAN switching is closely aligned with the behavior of traditional multiport bridges. The main difference between a LAN switch and a traditional bridge is that the LAN switch provides far greater port density. Because of the implementation of application-specific integrated circuit (ASIC) technology of the LAN switch, the cost per port is significantly less than a traditional bridge and the LAN switch makes an excellent replacement for shared-hub technology. With the implementation of VLANs, LAN switches also allow greater scalability than traditional bridges. A **LAN switch** can be defined as a device that acts at layer 2 of the OSI model and a **router** as a device that acts at layer 3 of the OSI model. It should be made clear, however, that the functions of a LAN switch and a router are merging; LAN switches are capable of some layer 3 functionality and routers do implement layer 2 functionality.

In the following sections switching and routing, including the services they provide, will be compared. It will become clear that a campus network needs both **switching (layer 2)** and **routing (layer 3) services**. The only way to achieve the appropriate mix of layer 2 and layer 3 services is to impose a multiVLAN architecture on the switched infrastructure.

A Comparison of Switching and Routing

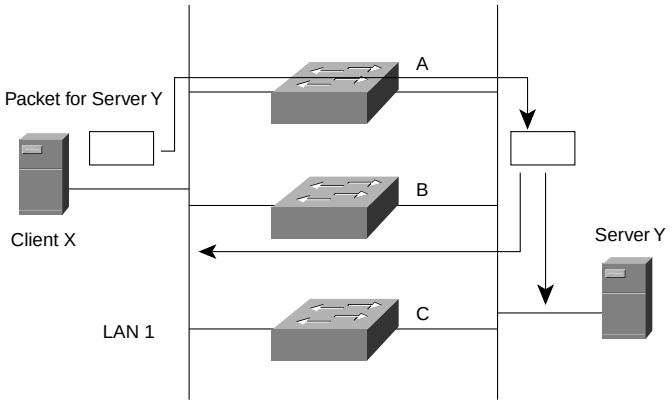
The following discussion provides an overview of the key aspects that should be considered when examining the relative merits of switching or routing.

Topology

Loops

LAN switches are susceptible to topology loops; an example of a possible looping situation is given in Figure 1.

Figure 1. A Potential Topology Loop.



In this scenario it is possible for packets from client X to get switched by A and then for B to put the same packet back onto LAN 1. In this situation not only do packets loop, they also undergo multiple replications. For this reason the 802.1d Spanning-Tree Protocol needs to be run in any potential looping topologies. The Spanning-Tree Protocol uses a conclusion from graph theory as a basis for constructing a topology that does not contain any loops. Graph theory states that:

> "For any connected graph consisting of nodes and edges connecting pairs of nodes, there is a spanning tree of edges that maintains the connectivity of the graph but contains no loops."

Consequently, only a subset of the network topology is used for forwarding data (certain links are put into blocking mode). The nature of routing allows loop freedom and the use of optimal paths.

Convergence

Transparent switching topology decisions are made locally because of the convergence of bridge protocol data units (BPDUs) that are exchanged between neighbor switches. Therefore, convergence can be adversely affected by the size of the switch domain. In a routed environment sophisticated routing protocols, such as Open Shortest Path First (OSPF) Version 2 (J. Moy, RFC 1583) or Enhanced Interior Gateway Routing Protocol (Enhanced IGRP) (Cisco Systems), maintain concurrent topological databases of the network and allow the network to converge quickly.

Broadcasts

LAN switches forward all broadcast, multicast, and unknown address frames out all ports. One must therefore be cautious in the design of a modern distributed system network where broadcast messages are used to resolve addresses and dynamically discover network resources, such as file servers. Broadcasts originating from each segment are received by every computer device in the entire switched internetwork. Most of these devices simply discard the messages as irrelevant, so large amounts of bandwidth may be wasted. In some cases the circulation of broadcast messages can become so severe that virtually no bandwidth remains for application data. In this case new network connections cannot be established and existing connections may be dropped. This situation is often referred to as a “broadcast storm.” The probability of broadcast storms increases as the switched internetwork grows. Routed networks, on the other hand, do not forward broadcast messages and, therefore, are NOT subject to broadcast storms.

Subnetworking

Transparently switched internetworks are composed of physically separate segments, but they are logically considered to be one large network (that is, one IP subnet). This behavior is inherent to the way that LAN switches work; that is, they operate at layer 2 of the OSI model and as such have to provide connectivity to end nodes as if each end node were on the same cable. Link layer (layer 2) addressing assumes a flat address space with universally unique addresses. As routers operate at layer 3 of the OSI model they are able to adhere to and formulate a hierarchical addressing structure. Hence a routed network is able to tie a logical addressing structure to a physical infrastructure by the use of, for example, TCP/IP subnets or IPX networks for each segment. Traffic flow in a switched (flat) network is therefore inherently different from traffic flow in a routed (hierarchical) network. Hierarchical networks offer more flexible traffic flow than flat networks as they are able to use the network hierarchy to determine optimal paths and contain broadcast domains.

Security

Routers and switches both have information available to them that can be used to create more secure networks. LAN switches may use custom filters to provide simple access control based upon destination address, source address, protocol type, packet length, and offset bits within the frame. Routers can filter on logical network addresses and provide control based upon options available in the network layer protocol. For example, specific TCP/IP socket information can be permitted or denied for a range of network addresses.

Media Dependence

Two factors need to be considered with regard to the media that an end node is connected to or the media interconnecting the network. Maximum frame size, often referred to as maximum transfer unit (MTU), differs for various network media. Table 1 illustrates how frame sizes can vary for different media.

Table 1. Media Frame Size Constraints

Media	Minimum Valid Frame	Maximum Valid Frame
Ethernet	64 Bytes	1518 Bytes
Token Ring	32 Bytes	16K Theoretical 4K Normal
FDDI	32 Bytes	4400 Bytes
ATM:		
• LAN E • Classical IP	• 64 Bytes • 64 Bytes	• 1518 Bytes • 9180 Bytes
Serial	14 Bytes HDLC	No limit 4.5K Usual

In effect, when dissimilar LANs are switched, the MTU used by end systems must be the lowest common denominator of all the switched LANs. This setup limits throughput and can seriously compromise performance over a relatively fast link such as Fiber Distributed Data Interface (FDDI) or Asynchronous Transfer Mode (ATM). However, because most network layer (layer 3) implementations can fragment and reassemble packets that are too large for a particular subnetwork, different M Us can be accommodated in routed networks. In this fashion throughput can be maximized in routed networks. A second factor that should be considered is that switching is datalink (layer 2) dependent and requires a translation function to switch between dissimilar media (that is, Ethernet to Token Ring or Ethernet to FDDI). This scenario can lead to nontrivial problems, such as noncanonical versus canonical Token Ring to Ethernet Media Access Control (MAC) format conversion. In contrast, routing is network layer (layer 3) protocol specific and operates essentially independent of physical media properties using an address resolution algorithm (that is, Novell node address = MAC address) or protocol (that is, Address Resolution Protocol [ARP]) mapping to resolve layer 2 to 3 address differences.

Benefits of Switching (Layer 2 Services)

Some of the features that LAN switches can potentially offer include:

- *Performance*

  LAN switches enable “microsegmentation,” which provides excellent performance for individual users. The extent of microsegmentation (whether a LAN switch has 4 to 5 users or a single user per switched port) determines the bandwidth available to individual users.

- *Switched Domains*

  This feature enables the grouping of individual ports into switched workgroups, restricting broadcast and multicast traffic to designated ports. It is commonly known as VLANs. Communication between switched domains requires a router.

- *Automated Packet Recognition and Translation*

  This feature affords the ability to translate as per the “IEEE 802.1h bridging to 802 LANs” mapping guidelines. (For example, Ethernet raw MAC frame format to FDDI Subnetwork Access Protocol [SNAP] automatically. The alternative is to send nonstandard FDDI raw, which most devices will not be able to recognize.) This method is still translation bridging but will automate the packet type translation to standards-based recommendations.

- *IP fragmentation (RFC791)*

  This feature addresses the varying MTU problems found in traditional bridges for IP packets. (It is not salient for IPX, which will negotiate maximum permissible MTU at client/server handshake.)

- *Broadcast Suppression Filters*

  This feature mitigates the effect of broadcast storms by allowing explicit filtering per port that dictates the amount of broadcast frames that are propagated over a chosen time interval. This solution is not a panacea to the adverse effects of broadcast storms. (For example, such features are unable to differentiate between broadcast messages that are used to resolve addresses and dynamically discover network resources such as file servers and undesirable broadcasts such as excessive unfiltered Service Advertisement Protocol [SAP] packets.)

- *Multicast Support for Port Membership Registration of Multicast Groups via Internet Group Management Protocol (IGMP)*

  When used in conjunction with multicast routing protocols (for example, Protocol Independent Multicast [PIM]), this feature affords an efficient distribution of multicast packets in a switched environment. The alternative is to multicast individual packets to each member port on the switch.

- *Subset of Layer 3 Routing Functions*

  For example, some LAN switches allow basic IP layer 3 routing. If IP routing is configured, the switch examines the protocol type of a received packet. If the protocol type is an IP packet, the IP routing code will handle the packet; if the protocol type is not an IP packet, the switch code will handle the packet. This feature allows network managers to permit multiple workgroups to use common switching resources without compromising network security and allows latitude on definition of common resources for multiple workgroups like file, mail, or print servers.

Benefits of Routing (Layer 3 Services)

In a typical campus network that supports multiple protocols such as IP, IPX, and AppleTalk, the role of the router is to prevent broadcasts from radiating out to all segments. However, the router is also responsible for providing services to each LAN segment. Some examples of services to the network that are provided by the router are:

- *IP*—proxy ARP, Internet Control Message Protocol (ICMP)
- *IPX*—SAP tables
- *AppleTalk*—Zone Information Protocol (ZIP) tables
- *Network Management*—Simple Network Management Protocol (SNMP) queries
- *DHCP*
- *IPX*—Get Nearest Server (GNS) queries
- *AppleTalk*—Name Binding Protocol (NBP)

It would be extremely burdensome for a single router to provide these services to hundreds of hosts on a flat virtual network. The router would get bombarded with a myriad of requests needing replies, severely taxing its processor. Therefore, some thought must be given to how many routers can provide reliable services to a given subset of VLANs. Some kind of hierarchical design would need to be considered.

Some of the benefits of routing include:

- *Broadcast Segmentation*

  Routers prevent broadcast traffic from traversing segments unnecessarily. Appendix B focuses on the effects of broadcast and multicast traffic on LAN hosts. Although VLANs reduce the effects of broadcast radiation, routers are essential for communication between VLANs.

- *Media Transition*

In the past, routers have traditionally been used to connect networks of different media types, taking care of the OSI layer 3 address translations and fragmentation requirements. This setup must still be used in new switched LAN designs. Most of the switching will be done within like media (such as Ethernet switches, Token Ring switches, and FDDI switches), with some capability of connecting to another media type. However, if a requirement for a campus switched network design is to provide high-speed connectivity between unlike media, routers need to play a significant part in the design.

- *Redundancy and Load Balancing*

  By virtue of the fact that routers run highly sophisticated routing protocols, routers are able to provide a high degree of both redundancy and load balancing.

- *Security*

  The fact that routers are essentially layer 3 devices means that routers can provide a high degree of packet manipulation in order to provide excellent security. Routers are able to restrict network layer addresses (access lists) and even look at particular application ports.

Summary

Switching provides the necessary bandwidth for today's new applications and helps to divorce the physical from the logical network. A particular subnet no longer needs to be locked into a physical segment, hence nodes located in several different locations (segments) can all be a part of the same logical subnetwork. To successfully achieve this concept of "virtual networking," a mixture of layer 2 and layer 3 services must be available.

Network Design Considerations

Several constraints govern campus network design when deploying LAN switching; some of the key points of each of these follow:

- Two questions drive the design of VLANs: "How homogeneous is the network?" and "How well-behaved are the VLANs?" Homogeneity refers to protocols (IP, IPX, or AppleTalk) or end stations (PCs or SUN workstations). The idea is to define VLANs along the lines of protocols or end stations and then design this VLAN topology within defined scalability constraints. Well-behaved means that 80 percent or more of the traffic is local to an individual VLAN. In an example where Marketing, MIS, and Engineering are all using individual VLAN segments, the 80-percent rule is violated. The Marketing employee reads mail from MIS, mounts servers from Engineering, and sends email to members of the Marketing VLAN. Clearly these networks do not lend themselves to clean segmentation, and detailed analysis of traffic patterns and data flows is necessary to design these VLANs well.
- Broadcast/multicast *background radiation*, as described in Appendix B, is a second constraint. Background radiation can become fatal; that is, 100 percent of the processors of a network can be consumed by broadcast and multicast activity. Guidelines for the number of nodes that should be in a single broadcast domain are given in the section entitled "Scalability."
- Bandwidth available to access layer 3 services and the location of layer 3 services within the hierarchical network model affect the design of campus networks.
- The final constraint governing campus network design is the concept of "administrative boundaries" and where they exist. Switching has the effect of "flattening" networks, and the deployment of switching by someone outside a user's administrative boundary could adversely affect the user's network.

This section will discuss what it means to design a good network, what the purpose of the three-layer hierarchical architecture is, and how campus network design has evolved. Within this framework three generic network topologies are presented. The ideal way to use this section of the paper is to examine the proposed topologies in conjunction with the scalability issues and the interoperability/availability matrices that highlight the features that are available on which platforms and when.

Good Network Design

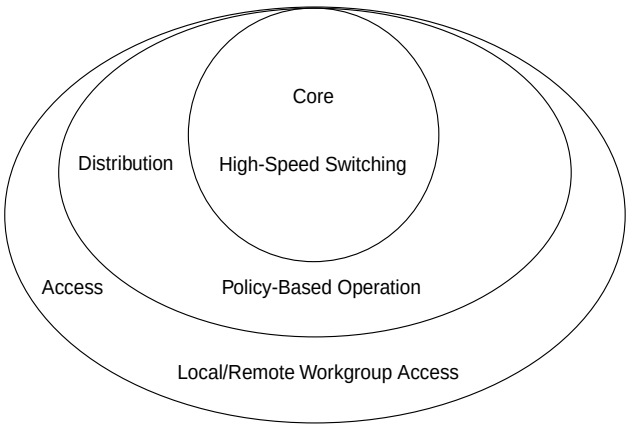
A well-designed network has to account for a multitude of facets. Some of the key network design principles follow:

- Networks should be built in a modular or hierarchical fashion. Maintaining this hierarchy allows for autonomous segments or areas of the network to be internetworked together.
- Traffic or data flows need to be characterized both for applications and protocols. Application data flow will profile client/server interaction and help with resource allocation (for instance, the number of clients using a particular server or the number of client workstations on a segment). Protocol behavior also must be analyzed carefully for each protocol. The section entitled “The Role of Routing in Switched Networks” examines the services provided by routers with respect to protocol control, address aggregation, and so on.
- Bandwidth availability must be analyzed; there should not be an order of magnitude difference between the different layers of the hierarchical model. It is important to remember that the hierarchical model refers to conceptual layers that provide functionality. The actual demarcation between layers does not necessarily have to be a physical link—it can also refer to the backplane of a particular device.
- Single points of failure must be examined carefully. Redundancy should exist in the network so that a single failure does not isolate any portion of the network. Two aspects of redundancy should be considered, backup and load balancing. In the event of a failure in the network, an alternate or backup path should be available; load balancing is exercised when two or more paths exist and the network has mechanisms to distribute network traffic across valid paths. The level of redundancy required in a particular network varies greatly from network to network.

Hierarchical Network Design

Figure 2 shows a high-level view of the various aspects of a hierarchical network design. A hierarchical network design presents three layers: the core, the distribution, and the access layers, and each of these provides a different function. The three layers do not need to exist in clear and distinct physical entities; they are defined to aid successful network design and as such represent **functionality** that needs to exist in a network. The occurrence of each layer can be in distinct routers or switches, represented by a physical media, combined in a single box, or a particular layer can be omitted altogether. For optimum performance, however, hierarchy should be maintained.

Figure 2. Hierarchical Network Structure



Core

The core layer is the high-speed switching backbone and should be designed to be “lean and mean.” In other words, the sole purpose of the core layer of the network is to switch packets as fast as possible. This layer of the network should not get involved in “expensive” packet manipulation, that is to say, anything that slows the switching of the packet down (access lists, filtering, and so on).

Distribution

The distribution layer of the network, the demarcation point between the access and core layers of the network, helps to define and differentiate the core. This layer provides boundary definition, and it is a location for the potentially “expensive” packet manipulations that preferably are avoided in the core. In the campus environment the distribution layer can represent a multitude of functions, some of which are:

- Address or area aggregation
- Departmental or workgroup access
- Broadcast/multicast domain definition
- InterVLAN routing
- Any media translations should happen at this layer
- Security

In the noncampus environment this layer can be a redistribution point between routing domains or the demarcation between static and dynamic routing protocols. It also should be the point at which remote sites access the corporate network. The distribution layer can be summarized as the layer that provides “policy-based connectivity.”

Access

The access layer of the network is the point at which end users are allowed into the network. This layer can provide further “tuning” in terms of filtering or access lists, but its key function is to provide access for end users into the network. In the campus environment some of the functions represented by this layer are:

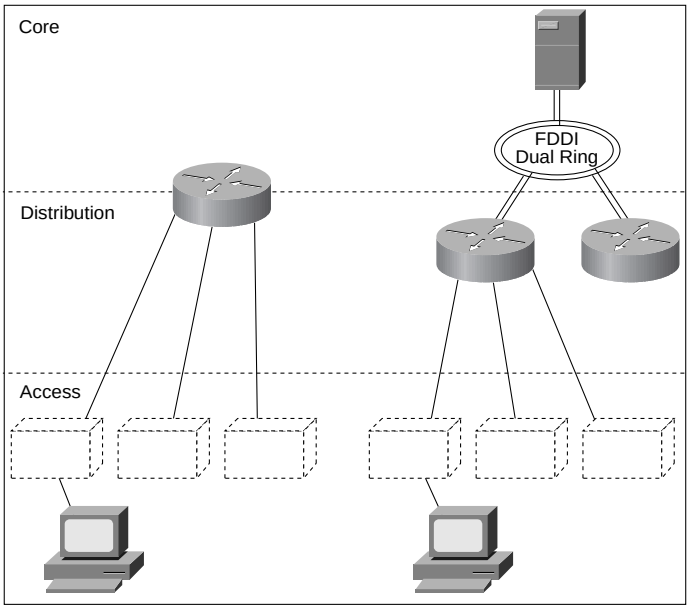
- Shared bandwidth
- Switched bandwidth
- MAC layer filtering (possibly)
- Microsegmentation

In the noncampus environment this layer provides access to remote sites into the corporate network via some wide-area technology (Frame Relay, Integrated Services Digital Network [ISDN], leased line).

Campus LAN Evolution

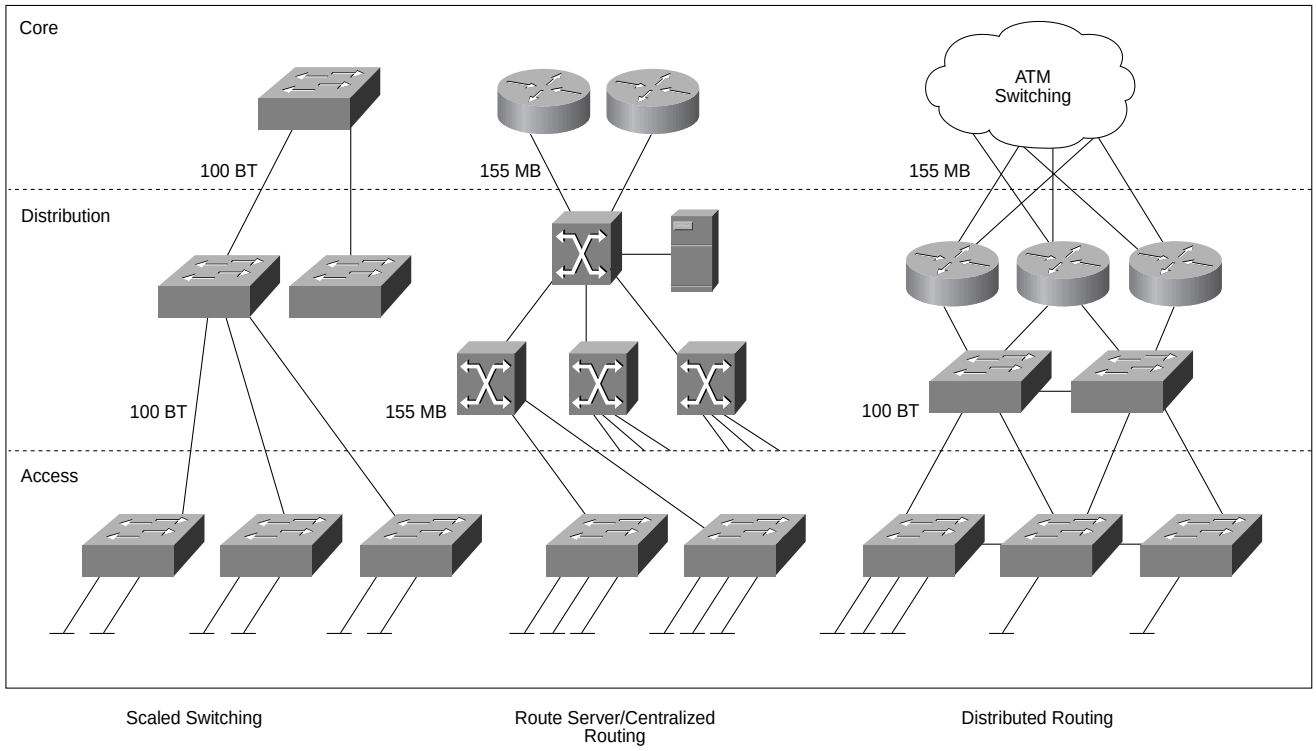
With respect to the hierarchical model, traditional campus LANs have followed one of two designs, either the “Backbone in a Box” or a “Distributed Backbone.” In the “Backbone in a Box,” the core and distribution layers are present in a single entity, the backbone router. Core functionality is represented by the backplane of the router and distribution by the interface cards of the router. Access for end users is via individual or chassis-based hubs. This design suffers from scalability constraints as the router can be in only one physical location and, therefore, everything needs to hub back to one location. Also all distribution functionality has to be accomplished by a single router, and this scenario can cause high CPU overload. The design can be scaled by using a high-speed backbone media, typically FDDI, to allow spreading of distribution functionality among several routers. This spreading also allows the distributed backbone to traverse floors in a building or even among buildings on a campus.

Figure 3. Traditional Campus Designs



Traditional campus networks are evolving quickly and switching is rapidly being deployed at all levels of the network, from the desktop to the backbone. Figure 4 represents three topologies that demonstrate the evolution of traditional campus networks. These topologies are proposed as possible generic network designs. The relative merits and constraints governing each of these topologies are presented in the sections that follow.

Figure 4. New Campus Designs



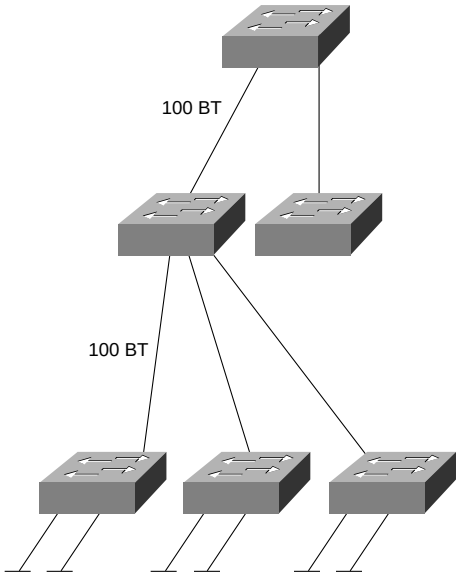
Proposed Topologies

The following topologies are proposed as generic network designs:

- Scaled switching
- Route server/centralized routing
- Distributed routing

Scaled Switching

Figure 5. Scaled Switching

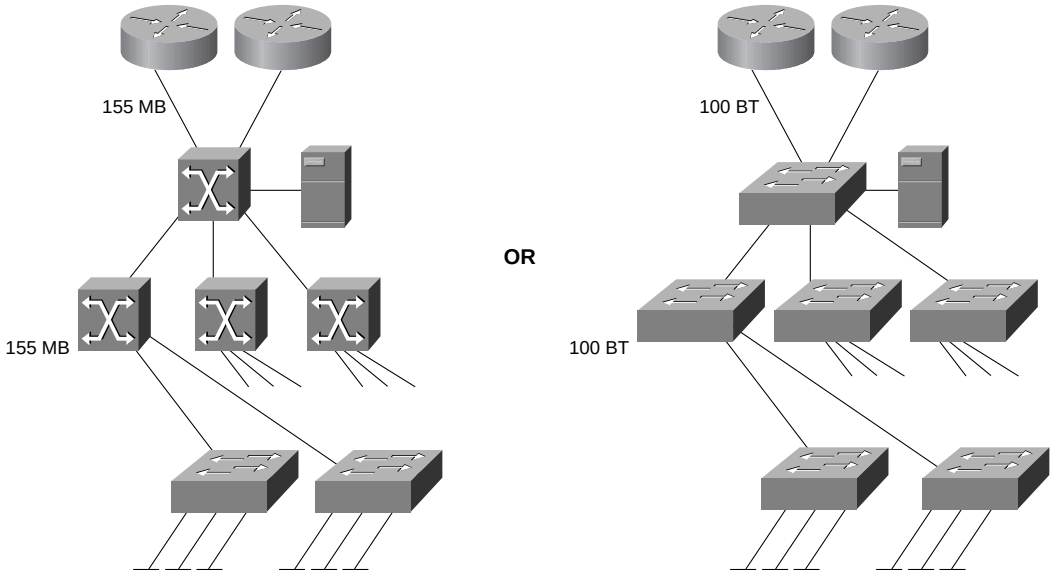


Scaled switching (Figure 5) deploys LAN switching at all levels of the network without the use of routers. This topology is economical and easy to install for a small campus network; it can be considered a “plug and play” network. It does not require knowledge of address structure, it is easy to manage, and all users can communicate with each other. However, since this network comprises a single broadcast domain, it MUST operate within the scalability rules defined in Appendix B. This topology is given only for completeness as it follows the hierarchical model only physically; it cannot implement the logical functions of the hierarchical model.

Key Issues

- Scalability limited by the number of hosts and type of applications running in the network
  - IP: 500 end stations (250 is a practical limit—class C address)
  - IPX: 300 end stations
  - AppleTalk: 200 end stations
- Possible to migrate to VLANs in order to limit the broadcast domains (however, no interVLAN communication is possible without the use of routers)

Figure 6. Route Server/Centralized Routing Setup



The route server/centralized routing topology deploys LAN switching at the access layer of the network, and either ATM switching or LAN switching at the distribution layer of the network. (See Figure 6.) Routers comprise the core layer of the network. The choice of architecture at the distribution layer of this topology, ATM, 100 BT switching, or FDDI switching, can be driven by factors such as cost or willingness to adopt leading-edge technology. This architecture has also been referred to as "Router on a Stick."

In the case of ATM in the distribution layer, LANE is used and some of the key issues that need to be considered are:

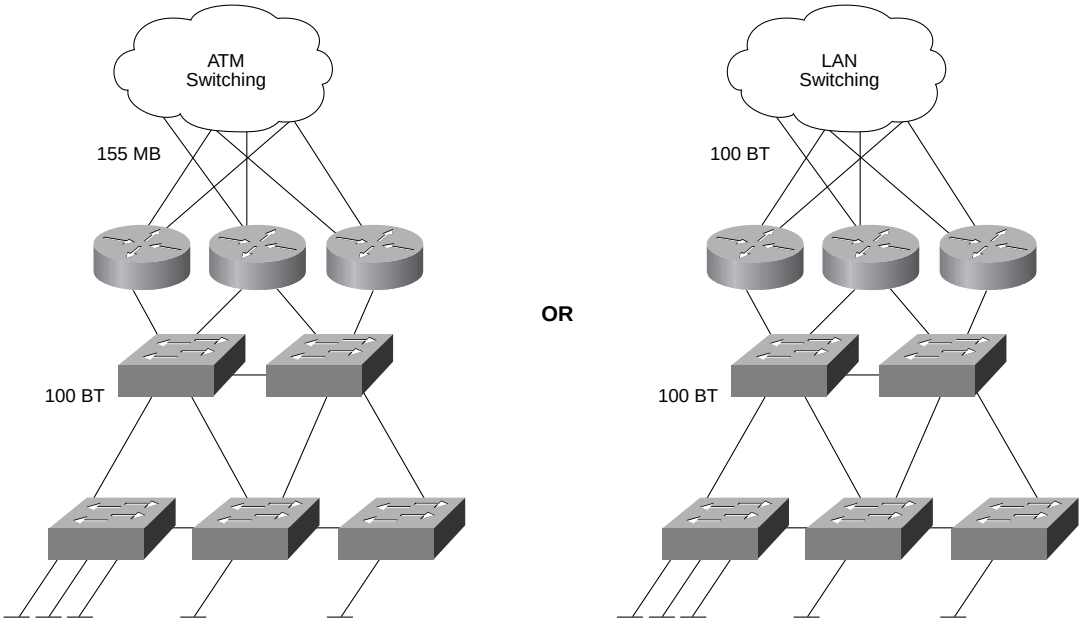
- LANE support on routers
- LANE support on the switch
- Support for UNI 3.X signaling (including point-to-multipoint signaling)
- Although it appears that there is redundancy in the distribution layer, it is possible that redundancy is provided by a virtual permanent virtual circuit/switched virtual circuit (PVC/SVC) mesh and there is actually a single point of failure in a single ATM switch
- Scalability (as detailed in the section entitled "Scalability")

In the case of LAN switching at the distribution layer some of the key issues are:

- Support of the appropriate VLAN muxing technology by the high-speed ports of the LAN switches being deployed
- With redundant LAN switches deployed at the distribution layer, Spanning-Tree Protocol needs to be run and some links will be blocked so that load balancing cannot occur
- Scalability (as detailed in the section entitled "Scalability")

This topology is attractive to customers for several reasons. In an environment where centralized servers or server farms are to be implemented, this topology is attractive as high-bandwidth, low-hop access can be provided to centralized resources. A second advantage of this topology is that it provides campus-wide VLANs. Although desirable, achieving these two goals adds cost. In this topology all layer 3 services are at the core layer of the network and the design is governed by the bandwidth available to access layer 3 services (that is, the 155- or 100-Mbps connections into the routers). This network design will scale well if VLANs are designed in such a way that the majority of resources are available locally (that is, in the VLAN). If VLANs can be designed such that 80 percent of traffic is intraVLAN and only 20 percent of traffic is interVLAN, the bandwidth available for access to layer 3 intelligence does not become an issue. However, if interVLAN traffic is high, then scalability is limited. In addition, since layer 3 services are not available to the access layers, this network will suffer the lack of layer 3 functionality (security, address aggregation, and so on).

Figure 7. Distributed Routing Topology



The distributed routing topology (Figure 7) deploys LAN switching at the access layer, routing at the distribution layer, and some high-speed switching mechanism at the core of the network. The core of the network is depicted in the figure by a cloud as it indicates a truly redundant switching fabric without a single point of failure. This network design follows the classic hierarchical network model both physically and logically. This network scales very well because it has high bandwidth available for access to layer 3 functionality. Some of the drawbacks with this design are that VLANs cannot exist across a campus; that is, the routers at the distribution layer provide a demarcation point between VLANs, and thus VLANs cannot be extended across the campus. If this technology were applied to a typical campus network, switches could be deployed on floors and “hubbed” back through switches to the routers in the basement. Therefore, this design allows for VLANs to exist between floors in a building but not between buildings. This topology can scale very well if, instead of using Fast Ethernet Interface Processors (FEIPs) to trunk VLANs into the routers, individual ports are used for each VLAN. The bottleneck to performance on the FEIP is dealing with ISL encapsulation/de-encapsulation; if this bottleneck is avoided and native traffic is switched by the FEIP, a vast increase in performance can be realized. The key issue that should be considered is scalability (as detailed in the section entitled “Scalability”).

VLAN Interoperability/Availability Matrix

Table 2 gives an overview of platforms and the technologies that each supports.

Table 2. Switch Platforms and VLAN Technologies Supported

	ISL	802.10	Lane
Catalyst 1200	N/A	Yes	N/A
Catalyst 1700/2000	No	No	Future
Catalyst 3000	Future	No	• ATM PVC (Q2 ’96) • ATM LANE (Q3 ’96)
Catalyst 5000	Yes	Yes	Yes
Routers	FEIP, Cisco IOS™ Release 11.1	All LANs, Cisco IOS Release 11.1	Yes, Cisco IOS Release 11.0
LightStream® 100	No	No	Yes, ILMI
LightStream 2020	No	Future	Q2 ’96
Zietnet	N/A	N/A	Yes

Switch Interoperability/Availability Matrix

Table 3 gives an overview of platforms and the technologies that each supports.

Table 3. Switch Platforms and Hardware Technologies Supported

	Fast Ethernet	FDDI	ATM
Catalyst 1200	No	Yes	No
Catalyst 1700/2000	Yes	Yes	Future
Catalyst 3000	Yes	No	Yes
Catalyst 5000	Yes	Yes	Yes
Routers	Yes, dependent on platform	Yes, dependent on platform	Yes, dependent on platform
LightStream 100	No	No	Yes
LightStream 2020	No	Yes	Yes
Zietnet	N/A	N/A	Yes

Scalability

Many facets of the network must be considered in order to determine network scalability. Numerous factors, such as bottlenecks in the network, resource allocation, CPU overhead, or available bandwidth, will affect the scalability of the network. Understanding the switching paths and consequently the throughput of the devices for various technologies is vital. The following guidelines are helpful:

- *VLAN Trunking Technologies*

  A clear understanding of the technologies (802.10, ISL, or LANE) that have been implemented and how they work is essential (as detailed in Appendix A). It enables recognition of the impact on the network in a worst case, a best case, or an average case scenario.

- *Throughput*

  Understanding the throughput (packets per second [pps]) of all devices in the network is also necessary. For switches, throughput should be at wire speed for all media, the only constraint being the aggregate throughput of the box. Routers exhibit vastly different throughputs, depending on the switching path that a particular packet follows.

- *Traffic Profiling*

  To understand "who is talking to what," network traffic profiles must be obtained, or, at a minimum, estimated. This information gives a picture of which users need which resources, what applications are running in the network, and where they are located. Traffic profiling is also necessary to design a virtual hierarchy or VLAN structure. A VLAN structure is imperative to give hierarchy and enable the benefits of layer 3 services in a "flat" layer 2 switched network. Many tools are available to help profile and understand the network. Some examples are:

  — Remote Monitoring (RMON) Analysis

  — Network Analyzers ("Top Talkers")

  — Network Management Tools

- *Broadcast/Multicast Applications*

  Applications such as video conferencing, distance-based learning (often referred to as multimedia), or systems for the stock exchange trading floors all make extensive use of broadcast/multicast traffic. This traffic can adversely affect the performance of ALL CPU in the network. Knowing which applications are being used and whether they will be generating broadcast or multicast traffic is desirable. The network should be designed so that these applications and their users do not affect other clients on the network. The Internetwork Design Guide (Cisco Connection Documentation CD-ROM) proposes design considerations and alternatives for a network that will be implementing multimedia or applications that are broadcast-intensive. The following guide for scaling the number of hosts in a particular broadcast domain (or VLAN) is approximate:

  — IP—500 end stations (250 is a practical limit—class C address)

  — IPX—300 end stations

  — AppleTalk—200 end stations

Requirements

- The following questions should be considered by a designer for a campus network:
- Size of community of interest? How many users?
- Geographic layout of the campus? How many buildings? Size: five miles? two miles? one building?
- Are users who share data located geographically?
- Cable plant: Do all cables run to central points?
- Do you spend a lot of time in adds, moves, and changes?
- What is the protocol mix? TCP/IP versus Novell versus NetBIOS versus Banyan VINES?
- How much traffic crosses the campus?
- Have you implemented server farms? If not, will you?
- Will you implement high-bandwidth applications, such as video conferencing?

Checklist

The following issues are relevant to the scalability of the network:

- How many ATM switched virtual circuits (SVCs) are supported by the various platforms?
  - c5000: 1024
  - c3000: 1900/card
  - 4500: 1024
  - 4700: 1024
  - 7000: 2048
  - 7500: 2048
  - LS2020: N/A PVCs only (4096/chassis)
  - LS100: 1024
- How many VLANs (ELANs) per device?

  The number is the **maximum** per chassis, an upper limit only; the actual number of VLANs that **should** be implemented is dependent on many factors.
  - c5000: 1024
  - c3000: 64 LECs/stack
  - 4500: 1024
  - 4700: 1024
  - 7000: 1024
  - 7500: 1024
  - LS2020: N/A
- How many LAN Emulation Configuration Servers (LECSs)/LAN Emulation Servers (LESs)/broadcast and unknown servers (BUSs) are supported per emulated LAN (ELAN)?

  Currently (as per the ATM Forum Specification), only one LECS per network and one LES/BUS per ELAN.
- Can RFC 1577 and LANE be run on the same interface?

  Yes. Once a LANE server is set up on a subinterface, the interface MTU is changed to 1516. Although RFC 1577 defaults to an MTU of 9180, not many users will actually use this because all end stations need to be manually changed to use the larger frame size. Thus both of these can be used on the same interface because an MTU of 1516 is acceptable within the RFC 1577 environment. The 7500 can support multiple MTUs on the same interface.
- What level of redundancy can Cisco support for VLANs?

  Redundant LECS/LES/BUS configurations are being investigated and will be incorporated into future LANE enhancements. The Spanning-Tree Protocol is supported.
- Is it possible to map ISL-to-802.10-to-LANE?

  Yes, it is possible to map between the different technologies.
- Does Hot Standby Router Protocol (HSRP) work in a VLAN environment to provide host redundancy?

  ISL uses a multicast address, VLAN information is carried in the ISL header, and the original packet is encapsulated within the ISL frame. Hence the multicast address used by HSRP is embedded in the ISL frame and is currently not looked at in an ISL environment. Thus it is impossible to use HSRP in an ISL environment. On the other hand, 802.10 and LANE do not encapsulate the original source and destination addresses, so HSRP will work in these two environments.

- Are all devices able to support LECS/LES/BUS functionality?

  The only device other than the routers that will support these functions is the C5000.

- How many broadcasts per second can the following devices handle?

  Broadcasts are process-switched, and the approximate rates are:

  — 4500: 8000–10000

  — 4700: 8000–10000

  — 7000: 1500–2000

  — 7500: 10,000–12,000

- What is the pps of the BUS on a per-device basis?

  The BUS is fast-switched. The fast switching rates are:

  — c5000: Will be wire speed

  — 4500: 30,000

  — 4700: 40,000

  — 7000: 20,000–25,000

  — 7500: 80,000–90,000

  — LS2020: N/A

- What is the pps of 802.10 traffic?

  802.10 traffic is fast-switched:

  — c5000: Not tested at the moment

  — 4500: 30,000

  — 4700: fast-switched40,000

  — 7000: 20,000–25,000

  — 7500: 80,000–90,000

  — LS2020: N/A

- What is the pps of ISL traffic?

  ISL traffic is fast-switched; however, there is an overhead for ISL encapsulations on the FEIP of the router:

  — c5000: Almost wire speed, minus the overhead for ISL encapsulation (30 bytes)

  — 4500: N/A

  — 4700: N/A

  — 7000: 10,000

  — 7500: 40,000

  — LS2020: N/A

Note: It is possible to greatly increase this throughput by using a single FEIP port into each VLAN (that is, avoid ISL encapsulation on the FEIP). In this case the fastest switching path, autonomous or fast switching (depending on protocol) on the 7000 and optimal or fast switching (depending on protocol) on the 7500, can be utilized.

Summary

Safe campus LAN designs use switches to replace traditional hubs and use an appropriate mix of routers to achieve minimal broadcast radiation. Three generic topologies have been presented; taken in conjunction with the interoperability/availability matrices and the scalability guidelines, they provide all the tools necessary to successfully design a switched campus network. A key choice has to be made by the customer with regard to campus-wide VLANs. The trade-off is that switching has to be deployed throughout the campus backbone or the core of the network, resulting in the loss of layer 3 services at the distribution layer.

Appendix A—Implementation of VLANs

This appendix describes the different methods of creating the logical groupings (or broadcast domains) that comprise various VLANs. With respect to this paper and the discussions herein, VLANs are differentiated by assigning each VLAN a "color" or VLAN identification (ID). ("Color" and "VLAN ID" are used interchangeably.) For example, engineering might all be in the "blue" VLAN, and manufacturing might all be in the "green" VLAN. The criteria for membership of a VLAN can be determined in three ways:

- *Port Basis*

  Each port on the router or switch can support only one color. This criterion is also referred to as a segment-based VLAN. All traffic within the VLAN is switched, and traffic between VLANs is routed (by an external router or a router within the switch).

- *Network Address Basis*

  Port-based VLANs are constrained by the fact that no layer 3 address recognition takes place, and consequently, IP, IPX, and AppleTalk networks all have to share the same VLAN definition. VLANs based on network addresses (layer 3 addresses) can differentiate between different protocols, and therefore can be defined on a per-protocol basis. This provision allows different virtual topologies for each protocol, each having its own set of rules, firewalls, and so on. Routing between VLANs occurs automatically, without the need for an external router or card. Thus there are potentially several colors on a port (setup also referred to as virtual subnet VLANs).

- *User-Defined Basis*

  This configuration typically gives the most flexibility, allowing VLANs to be defined and based on any field within a packet. VLANs can be defined on a protocol basis or be dependent on a particular IPX or NetBIOS service, for example. The simplest form of this type of VLAN is to group users based on MAC addresses.

Cisco's initial method of implementing VLANs will be on a per-port basis on both the Catalyst™ line and on routers. It should be noted that end station ports can support only a single VLAN, whereas trunking ports can support multiple VLANs. For management and efficient protocol operation, all nodes in a VLAN **should** be in the same subnet (for IP, IPX, or AppleTalk, and so on).

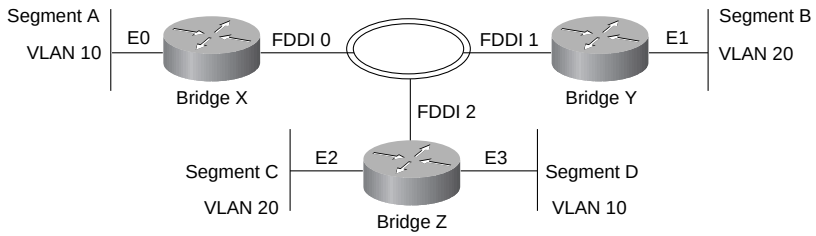
Cisco uses three types of technology to implement VLANs: the standards-based IEEE 802.10, the Cisco proprietary ISL, and LANE. They are all similar in that they are based on layer 2 bridge multiplexing mechanisms.

IEEE 802.10

The IEEE 802.10 standard provides a method for secure bridging of data across a shared metropolitan-area network (MAN) backbone. Cisco has initially implemented the relevant portions of the standard to allow the "coloring" of bridged traffic across high-speed backbones (FDDI, Fast Ethernet, Ethernet, Token Ring, serial links). Two strategies can be used to implement VLANs using 802.10, the only difference being whether the backbone is routed or bridged.

The first case has a bridged backbone and the objective is that bridged traffic goes only between segments A and D (both in VLAN 10) or between segments B and C (both in VLAN 20). (See Figure 8.)

Figure 8. 802.10-Based VLANs and a Switched Backbone

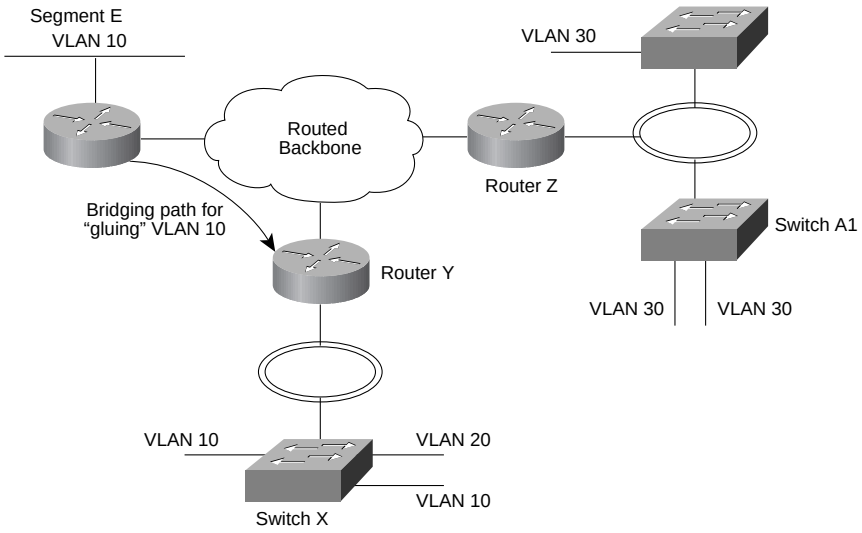


In this example, all Ethernet ports on bridges X, Y, and Z are in a VLAN and are to be VLAN interfaces. All FDDI interfaces in bridges X, Y, and Z are called transit bridge (TB) interfaces. To ensure that traffic from segment A destined for segment D on bridge Z is forwarded onto Ethernet 3 and not onto Ethernet 2, it is colored when it leaves bridge X. Bridge Z recognizes the color and knows that it must forward these frames onto Ethernet 3 and not onto Ethernet 2.

The "coloring" of traffic across the FDDI backbone is achieved by inserting a 16-byte header between the source address and the service access point (SAP) of frames leaving a bridge. This header contains a 4-byte VLAN ID or "color." A receiving bridge removes the header and forwards the intact frame to interfaces that match that VLAN color.

The second case has a routed backbone and the same constraints for VLANs 10 and 20.

Figure 9. 802.10-Based VLANs and a Raised Backbone



As stated earlier, it is important to be consistent with subnets and ensure that a single VLAN use only one subnet. In the example shown in Figure 9, VLAN 10 (subnet 10) is "split" and therefore has to be "glued" together by maintaining a bridged path for it through the network. If we consider switch X and a node in VLAN 20 (subnet 20), traffic is switched locally if appropriate. If traffic is destined for a node in VLAN 30 (subnet 30), router Y routes it through the backbone to router Z. This feature is possible because nodes in VLAN 20 (subnet 20) will always send any traffic that is destined for another subnet to its default gateway. This default gateway has to be the address of router Y on the FDDI ring. This example has used the term "default gateway," which is consistent with the IP model; however, other routable protocols behave in a similar manner (that is, they always send nonlocal traffic [off their subnet] to the local router).

To reiterate the key issues:

- VLANs must be consistent with the routed model (cannot split subnets)
- If subnets must be split, they must be "glued" together by a bridged path
- Normal routed behavior must be maintained for end nodes to correctly achieve routing between VLANs

- Networks must be designed carefully when integrating VLANs; the simplest choice would be not to split VLANs across a routed backbone

The difference between these two topologies is subtle. Some of the advantages/disadvantages for using one or the other are as follows.

Bridged Backbone

Advantages

- Propagates color information across entire network
- Allows greater scalability by extending bridge domains

Disadvantages

- Backbone runs bridging
- Broadcast scalability

Routed Backbone

Advantages

- No bridging on backbone; easy to integrate into existing internetwork; if subnets are split, they need to be glued together (that is, a bridged path has to be set up between switches)
- Can talk natively on the backbone

Disadvantages

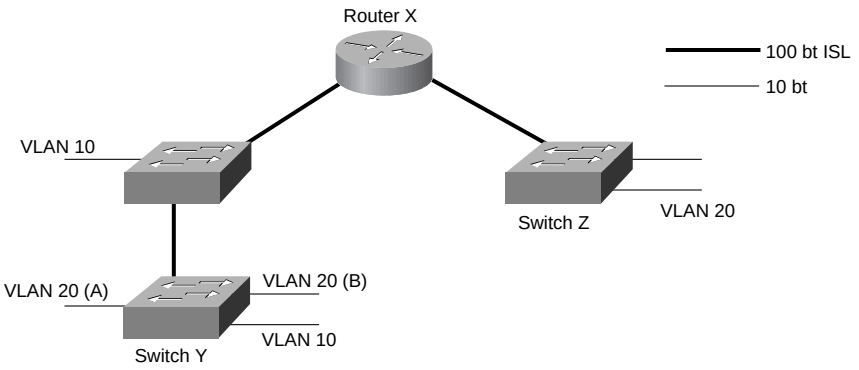
- Color information is not propagated across backbone, needs to be configured manually

Inter Switch Link

ISL is a Cisco proprietary protocol for interconnecting multiple switches and maintaining VLAN information as traffic goes between switches. This technology is very similar to IEEE 802.10 in that it is a method of multiplexing bridge groups over a high-speed backbone. It is defined only on Fast Ethernet. Many of the features described in the previous section also apply to ISL.

With ISL, an Ethernet frame is encapsulated within an ISL frame, and this maintains colors (VLAN IDs) between switches. ISL has a 30-byte header and contains a 2-byte VLAN ID. (See Figure 10.)

Figure 10. ISL-Based VLANs



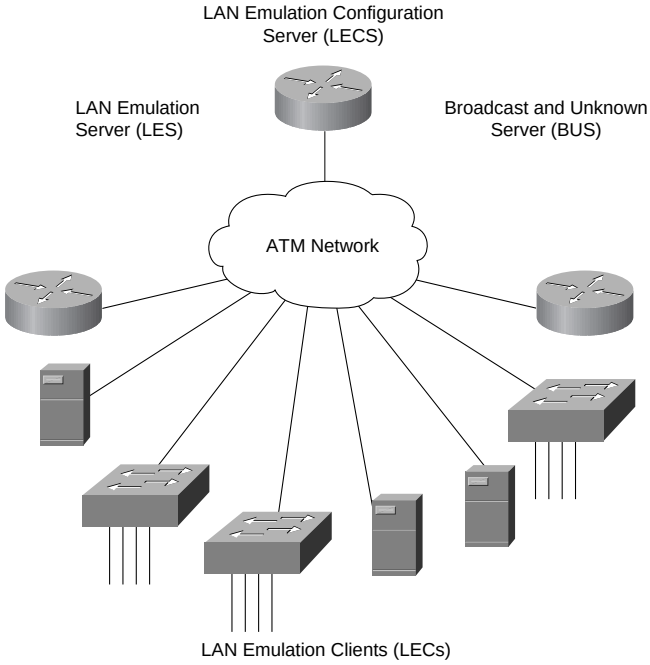
If we consider traffic for VLAN 20, switch Y switches traffic between segments A and B if appropriate. Otherwise it encapsulates traffic with an ISL header that identifies it as traffic for VLAN 20 and sends it through the interim switch to router X. Router X examines the ISL header and routes the packet to the appropriate interface, which could be through a routed network beyond router X (and is in this case) out the Fast Ethernet interface to switch Z. Switch Z receives the packet, removes the ISL header (noting that this packet is destined for VLAN 20), and switches it to all ports in VLAN 20 if this is a broadcast/multicast or to the appropriate port if a unicast.

To reiterate the key issues:

- VLANs must be consistent with the routed model (cannot split subnets)
- ISL is Cisco proprietary

LAN Emulation

Figure 11. Components of LANE



The ATM Forum has defined a standard for ATM LANE that provides stations that are attached via ATM the same capabilities that they are used to obtaining from legacy LANs. LANE is a service that provides easy interoperability between ATM-based workstations and devices connected to existing legacy LAN technology.

LANE is defined as a MAC encapsulation (layer 2), because this approach supports the largest number of existing layer 3 protocols. The end result is that all devices attached to an ELAN will appear to be on one bridged segment. In this way AppleTalk, IPX, and other protocols should have similar performance characteristics as in a traditional bridged environment.

The LANE components include the following:

- *LANE Client*

  End systems such as network interface card (NIC)-connected workstations, Catalyst 5000s, or Cisco 7x00s that support LANE require the implementation of a LANE client (LEC). The LEC emulates an interface to a legacy LAN to the higher level protocols. It performs data forwarding, address resolution, and registration of MAC addresses with the LES and communicates with other LECs via ATM virtual channel connections (VCCs).

- *LANE Server*

  The LES provides a central control point for all LECs. LECs maintain a Control Direct VCC to the LES to forward registration and control information. The LES maintains a point-to-multipoint VCC (known as the Control Distribute VCC) to all LECs, and only control information is forwarded on this VCC. As new LECs join the ATM ELAN, they are added as a leaf to the Control Distribute tree.

- *BUS*

  The BUS acts as a central point to distribute broadcasts and multicasts. ATM is essentially a point-to-point technology without “any-to-any” or “broadcast” support. LANE solved this problem by centralizing the broadcast support in the BUS. Each LEC must set up a Multicast Send VCC to the BUS. The BUS will then add the LEC as a leaf to its point-to-multipoint VCC, the Multicast Forward VCC.

  The BUS also acts as a multicast server. LANE is defined on ATM adaptation layer 5 (AAL5), which specifies a simple trailer to be appended to a frame before it is segmented into ATM cells. The problem is that there is no way to differentiate between ATM cells from different senders when multiplexed on a virtual channel. It is assumed that cells received will be in sequence, and when the end of message (EOM) cell arrives all the cells that have already arrived will be reassembled.

  The BUS must take the sequence of cells on each Multicast Send VCC and reassemble them into frames. When a full frame is received, it can be queued to send to all the LECs on the Multicast Forward VCC. This way all the cells from a particular data frame can be guaranteed to be sent in order and not interleaved with cells from any other data frames on the point-to-multipoint VCC.

- *LANE Configuration Server*

  This server maintains a database of LECs and the ELANs that they belong to. It accepts queries from LECs and responds with the appropriate VLAN identifier, namely the ATM address of the LES that serves the appropriate VLAN/ELAN. This database is maintained by the network administrator.

- It should be noted that, since LANE is defined at layer 2, the LECS is the only security checkpoint available. Once the LEC has been told where to find the LES, and has successfully joined the ELAN, the LEC is free to send any traffic (whether malicious or not) into the bridged ELAN. The only place for any layer 3 security filters is in the router that is routing this ELAN to other ELANs. Therefore, the larger the ELAN the greater exposure one must endure.

LANE Operation

In a typical LANE operation, first the LEC must find the LECS to discover which ELAN it should join. Specifically, it looks for the ATM address of the LES that serves the desired ELAN. To find the ATM address of the LECS, the LEC has three choices that it should attempt in the following order:

1. Query the ATM switch via Interim Local Management Interface (ILMI). The switch has a Management Information Base (MIB) variable set up with the ATM address of the LECS. The LEC can then contact the LECS with User-Network Interface (UNI) signaling.
2. Look for a fixed ATM address that is specified by the ATM Forum as the LECS ATM address.
3. Access permanent virtual circuit (PVC) 0/17, a “well-known PVC.”

The LEC creates a signaling packet with the ATM address of the LECS. It signals a Configure Direct VCC and then issues an LE_CONFIGURE_REQUEST on that VCC. The information in this request is compared with the data in the LECS database. The source ATM address is most commonly used to place a LEC into a specific ELAN. There may also be a default LES for those LECs that are not found in the database.

If a matching entry is found, a successful LE_CONFIGURE_RESPONSE is returned with the ATM address of the LES that serves the desired ELAN.

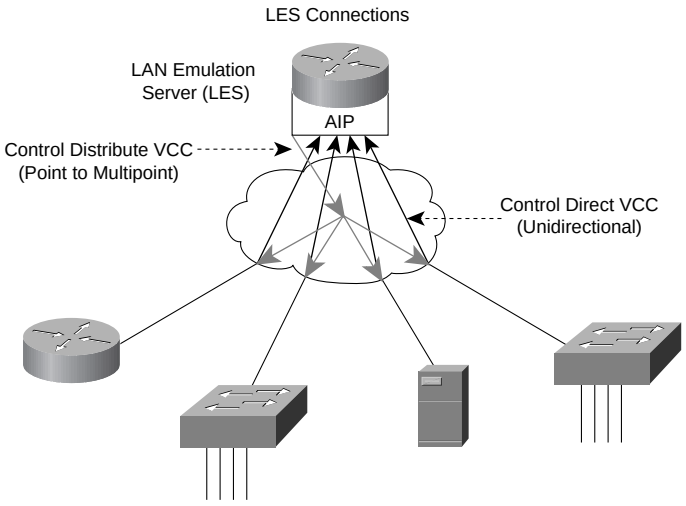
In Cisco’s first implementation, the LECS will reside on a subinterface of an ATM Interface Processor (AIP). It seems likely that this function will move to the network management station at some time in the future, because this is where all the VLAN membership information most likely resides.

Joining the LES

Once the LEC has discovered the ATM address of the desired LES, it drops the connection to the LECS, creates a signaling packet with the ATM address of the LES, and signals a Control Direct VCC. Upon successful VCC setup, the LES sends an LE_JOIN_REQUEST. This request contains the LEC ATM address as well as a MAC address that the LEC wants to register with the ELAN. This information is maintained so that no two LECs will register the same MAC or ATM addresses.

Upon receipt of the LE_JOIN_REQUEST, the LES checks with the LECS via its own open connection with the LECS and verifies the request, thus confirming the client's membership. Upon successful verification, the LES adds the LEC as a leaf of its point-to-multipoint Control Distribute VCC. Finally, the LES issues the LEC a successful LE_JOIN_RESPONSE that contains a LANE client ID (LECID), which is an identifier that is unique to the new client. This ID is used by the LEC to filter its own broadcasts from the BUS. (See Figure 12.)

Figure 12. LES Connections



Finding the BUS

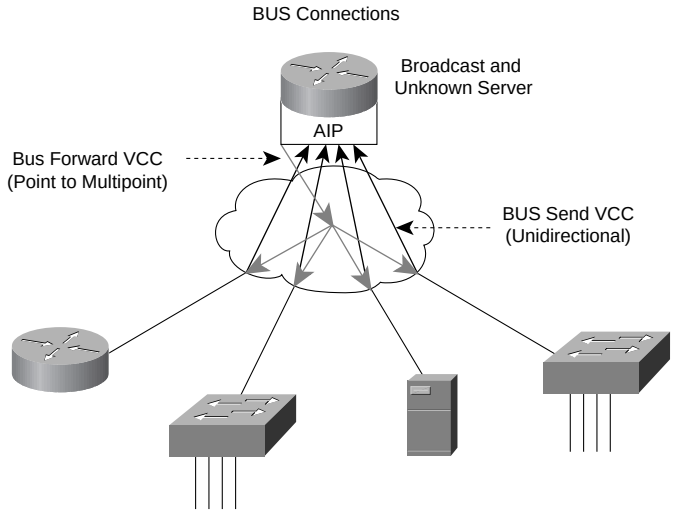
Once the LEC has successfully joined the LES, its first task is to find the ATM address of the BUS and join the broadcast group. The LEC creates an LE_ARP_REQUEST packet with the MAC address 0xFFFFFFFF. This special LE_ARP packet is sent on the Control Direct VCC to the LES. The LES recognizes that the LEC is looking for the BUS, responds with the ATM address of the BUS, and forwards that response on the Control Distribute VCC.

Joining the BUS

When the LEC has the ATM address of the BUS, its next action should be to create a signaling packet with that address and signal a Multicast Send VCC. Upon the receipt of the signaling request, the BUS adds the LEC as a leaf on its point-to-multipoint Multicast Forward VCC.

At this point the LEC has become a member of the ELAN. (See Figure 13.)

Figure 13. BUS Connections



LANE Operation

The real value of LANE is the ATM forwarding path for unicast traffic between LECs. When a LEC has a data packet to send to an unknown destination, it issues an LE_ARP_REQUEST to the LES on the Control Direct VCC. The LES forwards the request on the Control Distribute VCC so that all LEC stations hear it. In parallel, the unicast data packets are sent to the BUS, to be forwarded to all endpoints. This “flooding” is not the optimal path for unicast traffic, and this transmission path is rate-controlled to 10 packets per second (per the LANE standard). Unicast packets continue using the BUS until the LE_ARP has been resolved.

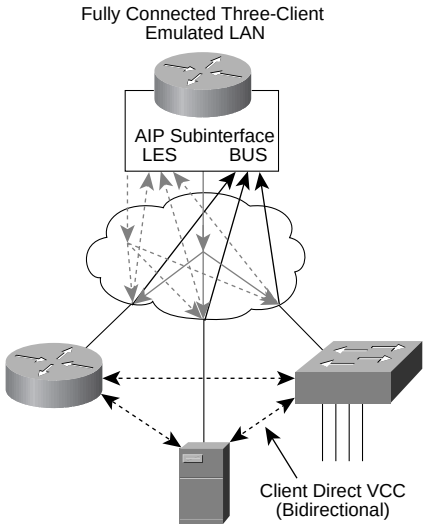
If there are bridge/switching devices with LEC software participating in the ELAN, they translate and forward the ARP on their LAN interfaces. Hopefully one of the LECs will issue an LE_ARP_RESPONSE and send it to the LES, which will forward it on the Control Distribute VCC so that all LECs can learn the new “MAC to ATM” address binding.

When the requesting LEC has received the LE_ARP_RESPONSE, it has the ATM address of the LEC that represents the MAC address being sought. The LEC should now signal the other LEC directly and set up a Data Direct VCC that will be used for unicast data between the LECs.

While waiting for the LE_ARP resolution, the LEC has been forwarding unicasts to the BUS, and now this new “optimal” path has become available. If the LEC switches immediately to the new path, it runs the risk of packets arriving out of order. The LANE standard has provided for a flush packet to guard against this situation.

When the Data Direct VCC becomes available, the LEC generates a flush packet and sends it to the BUS. When the LEC receives its own flush packet on the Multicast Forward VCC, it knows that all previously sent unicasts must have already been forwarded. It should now be safe to begin using the Data Direct VCC. (See Figure 14.)

Figure 14. Fully Connected Three-Client ELAN



Cisco's LANE Implementation

Cisco implemented LECS, LEC, LES, and BUS functionality in the AIP in Cisco IOS Release 11.0 These functions are defined on ATM subinterfaces (where one physical interface [such as an OC-3 fiber] is logically divided into up to 255 logical interfaces).

Some of the LANE design considerations are:

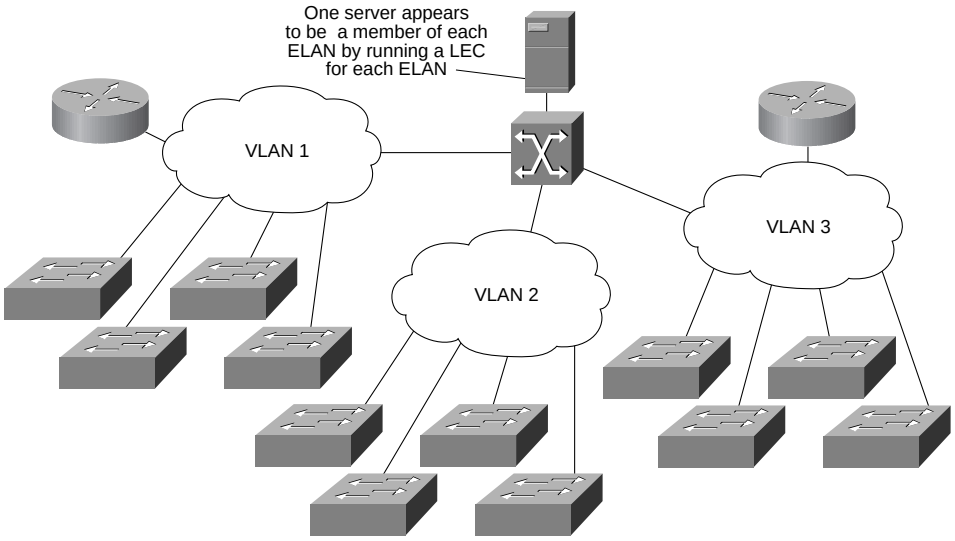
- One LECS supports all ELANs
- In each ELAN, there is one LES/BUS pair and some number of LECs
- The LES and BUS functionality must be defined on the same subinterface and cannot be separated
- There can be only one LES/BUS pair per subinterface
- There can be only one LES/BUS pair per ELAN
- The current LANE phase 1 standard *does not provide* for any LES/BUS redundancy
- The LECS and LES/BUS can be on different routers/bridges/workstations
- VCCs can be either SVCs or PVCs, although PVC design configuration and complexity would probably make anything more than a very small network prohibitively unmanageable and complex
- When defining VLANs with the Catalyst 5000, each VLAN color must be assigned to a different ELAN. The LES/BUS for each of these ELANs can reside on:
  - Different subinterfaces on the same AIP
  - Different AIPs in the same router
  - Different AIPs in different routers
- There can be only one LEC per subinterface. If a LEC and a LES/BUS pair share a subinterface, they are (by definition) in the same ELAN
- If a LEC on a router subinterface is assigned an IP, IPX, or AppleTalk address, that protocol is **routable** over that LEC. If there were multiple LECs present on a router and they were assigned protocol addresses, routing would occur between the ELANs. It is necessary to maintain subnet integrity (that is, an ELAN should be in only one subnet for a particular protocol) for routing between ELANs to function correctly.

Virtual Multihomed Servers

Several well-known servers, such as email and corporate servers, that almost everyone in an enterprise needs to access, usually exist in traditional networks. If these servers were located in only one virtual LAN, the benefits of VLANs would be lost with all the different workgroups being forced to *route* to access this common information source.

This problem can be solved with LANE and virtual multihomed servers. (See Figure 15.) NICs such as the Zeitnet allow the workstation/server to join up to eight different VLANs. This means that the server will appear in eight different ELANs, and to other members of that ELAN, the server would appear just like any other member. This feature greatly increases the performance of the network as a whole, because this common information is available directly through the optimal Data Direct VCC and does not have to be routed.

Figure 15. Virtual Multihomed Server



To multihome servers in the nonATM environment there are two possible choices:

- Have servers with multiple NICs (different connections into each VLAN)
- Have servers with NICs that understand the multiplexing technology being used on a backbone (that is, 802.10 or ISL)

Appendix B—Effects of Broadcast and Multicast Traffic

To a LAN host, a layer 2 LAN switch provides the same basic features as a bridge: LAN contention, or more specifically, the Ethernet collision domain, is reduced; “local” traffic is not forwarded to other segments; and configuration of the device is minimal. However, the problem with layer 2 switching is that, by definition, it must forward broadcast and multicast traffic to every segment. LAN or desktop protocols require these all-destination and all-vendor ID packets to be flooded throughout the subnet/broadcast domain, so that features like service advertisement and connection requests work properly. Routing technology and future technologies such as multilayer switching will constrain multicast and broadcast traffic through network-layer addressing. In this appendix, these OSI layer 2 and 3 issues and the effect this traffic has on LAN hosts and their CPU utilization will be examined.

All desktop protocols like AppleTalk, IPX, and IP need to efficiently communicate with the entire flat network or portions of it on a regular basis. At the MAC layer this communication is accomplished through “broadcasting” (sending packets to the 0xFFFFFFFF destination) and group broadcasting (“multicasting”), where the leading bit of vendor ID is set to 1. Every host with the specified vendor ID will process all multicasts to its group address.

The NIC on most hosts performs a filtering function that will send datagrams further up the protocol stack only if they are specifically destined to its MAC, broadcast, or multicast address. A broadcast packet passes through the unicast filtering process of the NIC and consequently affects the protocol stack of every host on the network. Generally this effect is evidenced by an

interrupt request to the main CPU caused by a signal from the host's NIC that a packet needs further processing. Some NICs provide the ability to select in hardware the multicast groups to listen to. Hence these cards may be more tolerant in a multicast-intensive environment but still must contend with broadcasts.

The term *broadcast radiation* refers to the broadcast or multicast traffic on a network segment that a given host must process even though that host derives no benefit from the data. A bridge, or switch, must forward broadcast and multicast packets throughout the network for the various functions of the LAN protocols to work properly. Even though unicast traffic is limited, broadcast radiation is uniform throughout a flat network. Thus a network manager must make a trade-off between ease of administration (of a flat network) and the amount of CPU performance degradation users are willing to tolerate.

This discussion is not to say that flat switched networks are bad, just that there is an upper limit to the size they can reach before broadcast radiation affects the hosts on the network. A switched network has comparable levels of broadcasts and multicasts to a hub or repeated network, but vastly more available bandwidth. A switch splits up collision domains, reducing collisions and increasing the utilization of available bandwidth as well as limiting unicast traffic. If a port is dedicated to a single host and its MAC address is cached, the host will see only unicast traffic that is specifically destined to it.

All protocols use the layer 2 broadcast or multicast function to varying degrees. Network layer protocols like Novell IPX, for example, periodically broadcast services out to the network, whereas IP sends broadcast requests (ARP discovery) for specific host destination addresses. IP uses ARP either the first time a host wants to communicate or after a host's ARP cache entry has aged out. What these protocols have in common is that different logically configured network segments are connected through routers, which eliminate network-wide broadcast propagation or broadcast radiation.

IP Scalability Issues

For IP and other layer 3 protocol-based networks, logical address configuration is a scalability mechanism. These nets tend to be limited in size to the address range available. Addressing domains provide natural splits, because IP nets are usually not bridged but routed together. In particular, IP is not a "chatty" protocol. The only advertisements that occur are routing updates being exchanged between routers.

In the switched environment, the IP protocol is of interest to network administrators who want to set up addressing schemes with a large address space. This could be a class B address with no or very few subnet bits, or a class A address with the standard eight-bit subnet mask. In this environment, IP would probably scale quite well. However, common IP clients like Sun SPARCstations still use interrupts to deal with broadcasts and multicasts. These stations would have performance scaling problems similar to those of the PC and MAC.

A SPARC 2 with a standard built-in Ethernet card can be effectively shut down by broadcasts flooding the network. While extreme, broadcast peaks of thousands of broadcasts per second have been observed in networks experiencing "broadcast storms." Testing in a controlled environment with a range of broadcasts and multicasts on the network shows an overall system degradation with as few as 100 broadcast/multicast frames per second.

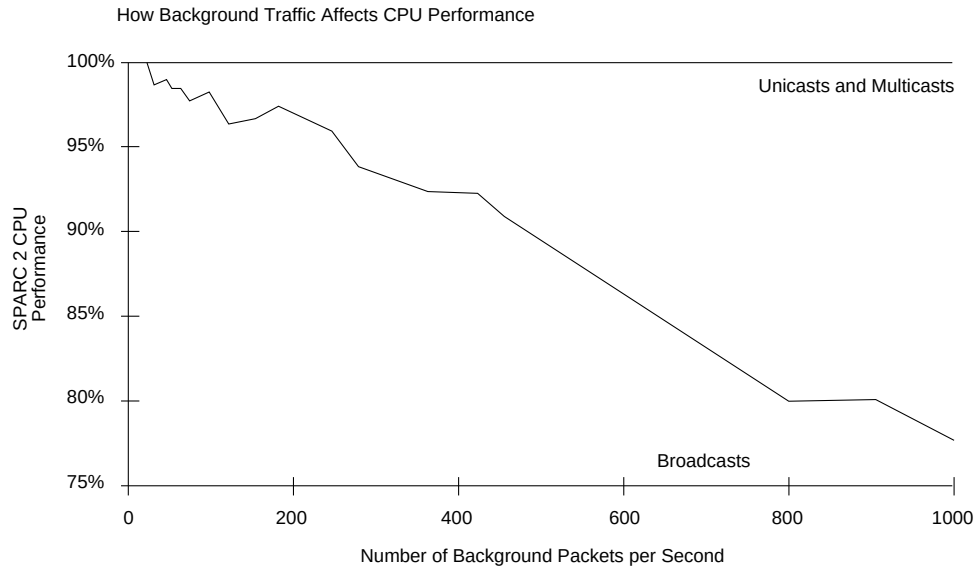
IP is now popular on many PCs and Macintoshes. Networks that include a variety of host types would have to be designed to work with the lowest common denominator. If Mac users complain about system performance until they get a SPARCstation or high-end station, over time the network will no longer be heterogeneous. In the end, the user will have spent a lot of money upgrading CPUs instead of the network.

Table 4 and Figure 16 give results of testing a generic SPARC 2 for the effect of broadcast/multicast traffic on performance. Testing was done by comparing CPU performance measurements with and without network loads.

Table 4. How Background Traffic Affects CPU Performance

Number of Broadcasts Received	Percent of CPU Consumed per Second
100	2.97
1000	22.55
2000	44.55
3000	68.55
3800	100.00

Figure 16. CPU Performance Degradation



Note: The SPARC 2 ran the SunOS Version 4.1.3 without an IP multicast kernel. Solaris 2.x automatically enables the IP multicast feature, which would allow multicast packets to affect the CPU.

IP broadcasts and multicasts are produced by two sources, workstations and routers. Routers in this case involve any router or workstation running the Routing Information Protocol (RIP) routing protocol. Many people configure all workstations to run RIP as a redundancy and reachability policy.

An IP workstation produces an ARP broadcast every time it needs to find a new IP address location on the network. Therefore, the command “telnet mumble.com” translates into an IP address through a Domain Naming System (DNS) search, then an ARP is broadcast to find the actual station. Generally, IP workstations keep 10 to 100 addresses for about two hours, so a typical ARP rate of a workstation might be about 50 addresses every two hours, or 0.007 ARPs per second. Thus, 2000 IP end stations would produce about 14 ARPs per second.

IP routers, generally running RIP, produce a string of packets about every 30 seconds, depending on the size of the routing table. The entire IP routing table is retransmitted every 30 seconds by each player. Therefore, if all 2000 workstations were configured for ROUTED (the RIP daemon), and each produced 50 packets every 30 seconds, 3333 broadcasts per second would be generated. At this point, no SPARCstation would be performing any work. Most network administrators configure a small number of routers, 5 to 10 usually, to speak RIP. For 10 RIP speakers and 50 RIP packets, about 16 broadcasts per second would be transmitted to all end stations in the switched LAN, or VLAN segment.

The effect of this scalability was investigated on a few networks, and results are given in Table 5.

Table 5. Average Broadcasts/Multicasts for Well-Behaved IP Networks

Number of End Stations	Average Percent CPU Loss for All Connected Stations
100	0.14
1,000	0.96
10,000	9.15

Although these numbers appear low, they represent an average, well-designed IP network that is not running the RIP routing protocol. When broadcast and multicast traffic peak because of “storm” behavior, peak CPU loss can be orders of magnitude greater than average. Broadcast “storms” can be caused by a device requesting information from a network that has grown too large. So many responses are sent to the original request that the device cannot process them, or the first request triggers similar requests from other devices that effectively block normal traffic flow on the network.

IP Multicasting

IP multicasting can adversely affect the performance of large, scaled switched networks. Although multicasting is an efficient way to send a multimedia/video stream to many users on a shared-media hub, it affects every user on a flat switched network. Networks that use multicasting applications should use smaller flat segments or VLANs to break up existing flat segments.

In addition, these applications should be migrated or upgraded to use the IGMP and PIM protocols. These new protocols allow a multicasting application to “negotiate” with routers, switches, and clients to determine which devices will be part of the multicast group. This negotiation helps limit the scope and impact of the multicast stream on the network as a whole.

High-quality multimedia video streams can be very bandwidth-intensive. A packet video application is available that can generate a seven-MB stream of multicast data that is sent to every segment in a switched network. Use of this application, however, can generate severe congestion. With the growing availability of these multicast-intensive applications, to protect a network it may be wise to make the “flat” portions of a network relatively small and move routers as close to the users as possible, making use of the multicast filtering capabilities inherent to routers through protocols such as IGMP and PIM.

IPX Scalability Issues

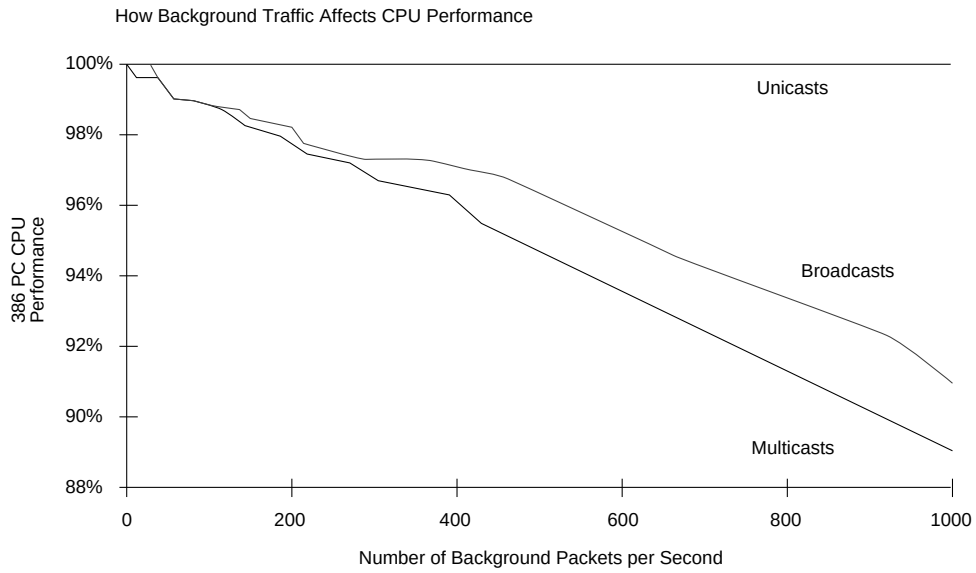
Many of today’s PC-based LANs use Novell servers and the Novell network operating system (NOS). In LANs that have a variety of host types, where the PCs are involved, the Novell NOS appears to be the client/server application of choice. This choice poses a few unique scaling problems:

- Every Novell server uses broadcast packets to identify itself and advertise its services and routes to other networks (RIP and SAP/Network Control Program [NCP]).
- Each client uses broadcasts to find the nearest server and connect to it.
- New (Version 4.0) Novell SNMP-based network management applications like NetExplorer periodically broadcast (IP/ARP) packets to discover changes in the network.

On a single-server network with one shared volume and no printer services, this scenario translates into one broadcast packet every four seconds. This amount might not seem like much, but the result would be just one server in an idle state. A large LAN with high-end servers might have up to 150 users per PC server. If the LAN has 900 users with a reasonably even distribution, that would optimally add up to six to seven servers. In an idle state with multiple shared volumes and printers, the average would be four broadcasts per second uniformly distributed. In a busy network where route and service requests are made frequently, this idle rate would peak significantly higher, up to 15 to 20 per second.

Figure 17 illustrates the effect that this background broadcast traffic would have on a client workstation. The performance of the client PC CPU (80386 at 25 MHz) was measured with the Norton Utilities System Information utility. Background traffic was generated with a Network General Sniffer and consisted of a broadcast destination packet and a multicast destination packet, with data of all zeroes. The client network interface drivers did not care whether the packets were Novell packet types or any other protocol. Every broadcast or multicast destination packet received by the workstation caused a CPU interrupt that reduced overall system performance. CPU performance was measurably affected by as few as 30 broadcast/multicast packets per second. Multicast packets had a slightly worse effect than broadcast packets.

Figure 17. How Background Traffic Affects CPU Performance



The scaling limitations of IPX include clients, servers, and routers. Clients use broadcasts to find local services, and because a large, scaled switched network is a single LAN entity, all servers respond with a multicast listing services they offer. Routers collect services not local to this switched entity and send out periodic SAP transmissions to describe the services offered on the entire network. Each router sends out one frame for every seven services on the network.

The effect of this scalability has been investigated on a few networks and the results are given in Table 6.

Table 6. Average Broadcasts/Multicasts for Well-Behaved IPX Networks

Number of End Stations	Average Percent of CPU Loss for All Connected Stations
100	0.12
1,000	0.22
10,000	3.15

These numbers represent a multihour, average operation. Peak traffic load and CPU loss per workstation can be orders of magnitude greater than with average traffic loads. A common scenario is that at 9 a.m. on Monday, everyone comes in and boots up their computers. Normally, in circumstances with an average level utilization or demand, the network can handle a reasonable number of stations. However, in circumstances when everyone requires service at once (a demand peak), the available network capacity can support a much lower number of stations. In determining network capacity requirements, peak demand levels and duration can be more important than average serviceability requirements.

AppleTalk Scalability Issues

AppleTalk uses multicasting extensively to advertise/request services and to resolve addresses. On startup, an AppleTalk host transmits a series of at least 20 packets aimed at resolving its network address (layer 3 AppleTalk node number) and local “zone” information. Except for the first packet, which is addressed to itself, these functions are resolved through AppleTalk multicasts.

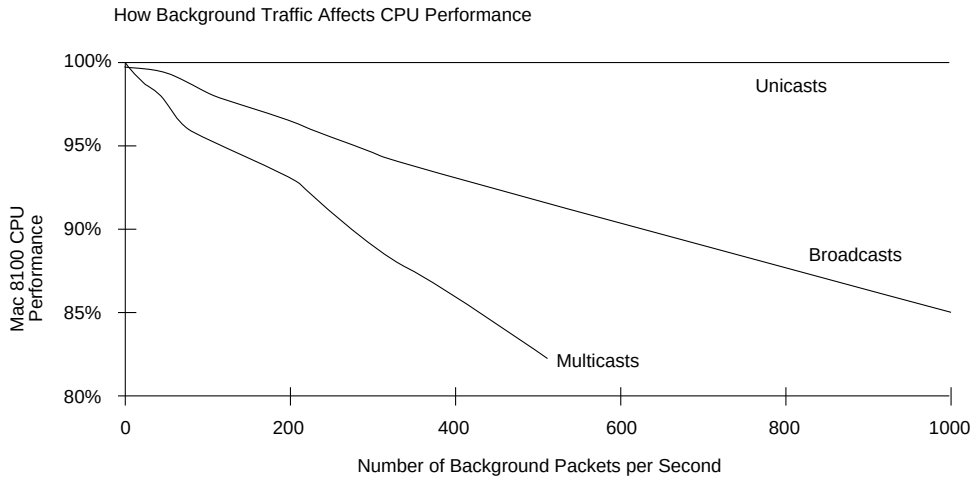
Two ill-behaved (in terms of overall network traffic) applications are the AppleTalk Chooser and a system Init called AutoRemounter. These applications are part of the Macintosh operating system. The Chooser is the software interface that helps the user select shared network services. It uses AppleTalk multicasts to find file servers, printers, and other services. If this application is left “open” by the user, it continues to send out network requests. When the user opens the Chooser and selects a type of service (printer for example), the Chooser transmits 45 multicasts at a rate of 1 packet per second (pps). If left open,

the Chooser sends five-packet bursts at the same rate with a progressively longer delay. If left open for several minutes, the Chooser reaches its maximum delay and transmits a five-packet burst every 270 seconds. By itself, this scenario does not pose a problem, but in a large network, these packets add to the overhead or background radiation traffic that a host must interpret and then discard.

The effects of broadcast radiation on a Macintosh CPU are illustrated in Figure 18. Two Macs were tested with similar results, an 8100 Power PC and a Mac IIci. Both CPUs were affected by as few as 15 broadcast or multicast frames per second. The AppleTalk stack performed better than the Novell PC stack in that the Mac did not suffer performance degradation with nonAppleTalk multicasts. The performance loss scaled linearly with the amount of background traffic.

Data collection on large “live,” flat AppleTalk networks showed a few broadcasts due to Mac Transmission Control Protocol (TCP), but significant bursts (up to 40 pps) of AppleTalk multicasts.

Figure 18. How Background Traffic Affects CPU Performance



AppleTalk has numerous multicast operations, from simple client discovery (the Chooser), which begins its discovery by broadcasting every 10 seconds, to network-wide operations such as Name Binding Protocol, which binds a client to a server, and Router Discovery Protocol, a RIP implementation that is transmitted by all routers and listened to by each station.

One major issue with large-scale AppleTalk networks is the slow LocalTalk-to-Ethernet connection device. This device fails in large AppleTalk networks because it has a limited ARP cache and can only process a few broadcasts per second. In the testing, a major broadcast storm arose because these devices lost their ability to receive Routing Table Maintenance Protocol (RTMP) updates. Once a storm arises, these devices begin ARPing for all known devices, thereby accelerating the network degradation because they cause their neighbor devices to fail and reARP.

The effect of this scalability has been investigated on a few networks and the results are in Table 7. The numbers in the table represent a multihour operation; peak CPU loss can be orders of magnitude greater than average. AppleTalk switched networks deteriorate as they are scaled. Consequently, the scope of these networks must be limited.

Table 7. Average and Peak Broadcasts/Multicasts for AppleTalk Networks

Number of End Stations	Percent of CPU Loss for All Connected Stations
Average	
• 100	• 0.28
• 1,000	• 02.10
• 10,000	• 16.94
Peak	
• 100	• 6
• 1000	• 58
• 10,000	• 100

Multiprotocol Effects

Research has indicated that the following protocols interact in a local switched environment:

- AppleTalk does not listen to any other layer 3 protocol
- IPX clients listen to IPX, AppleTalk, and IP
- IP clients listen to IP, IPX, and AppleTalk

The results imply that, if an AppleTalk network emits a lot of broadcast activity, it has a cumulative effect on IPX and IP.

Summary

Successful network designs contain a mix of appropriately scaled switching and routing. Switching, either as a global construction or a VLAN group, can be successful if well managed. Well-managed switches must include broadcast and multicast management. Within these constraints, the following guide should be used for layer 3 scaling:

- *IP*—500 end stations (250 is a practical limit—class C address)
- *IPX*—300 end stations
- *AppleTalk*—200 end stations

Mixed networks should probably not have more than 200 end stations.

Appendix C—Network Management

Two problems need to be addressed with VLAN management software. First, the administration and management of VLANs would be extremely difficult without a graphical user interface (GUI)-based network management tool. Second, the fact that VLANs allow membership regardless of physical location means that various VLANs (colors) can exist throughout the network infrastructure. The monitoring of traffic flow or "color monitoring" is pivotal in the maintenance of a healthy VLAN network. VlanDirector™ and CiscoFusion™ Manager are VLAN management software that provide this capability.

VlanDirector

VlanDirector is a VLAN configuration and management tool that is currently shipping and provides management support for the Catalyst 5000s connected via Fast Ethernet/Ethernet. The next release of VlanDirector will include product support for the Catalyst 3000 and 1200, including configuration and management of VLANs across FDDI rings. VlanDirector currently runs as a standalone application and comes bundled with CiscoView™. Supported platforms include Sun SPARCstations with Sun OS 4.1.3/4 and HP Unix versions with HP/UX. VlanDirector will also operate on top of SunNet Manager and HP OpenView (for Unix). VlanDirector will be available on IBM RS6000 workstations in Q1'96. VlanDirector is not integrated with CiscoWorks. A PC-based demonstration diskette is available that explains the functionality of the application (literature distribution #73501). The important features of VlanDirector include:

- Autotopology discovery via Cisco Discovery Protocol (CDP)
- A simplified VLAN naming window and associated directories offer flexibility and easy-to-use search functions for creating and changing VLAN names
- Represents the logical overlay of VLAN configurations on top of the physical topology; this feature is critical when designing and monitoring VLANs
- Logical views of configured VLANs include switch, link, and port membership windows, enabling users to audit and verify membership status
- Network managers can fine-tune the configurations of VLANs across interswitch links through simple drag and drop operations as well as by selecting multiple paths, choosing specific paths based on preferences, and manually adding and deleting paths
- A drag-and-drop VLAN configuration per switch port via CiscoView
- VLAN membership report generation
- Provides error reports when links between switches have been incorrectly configured
- Provides name search functions for locating and monitoring specific VLANs
- User-selectable color options concurrently display multiple configured VLANs, making visualizing and identifying VLAN configurations easier

In addition to the VlanDirector features, TrafficDirector™ will be available for RMON on switched ports and across VLANs. Examples of VlanDirector user interfaces are shown in Figures 19 through 21.

Figure 19. VlanDirector Directory Structure

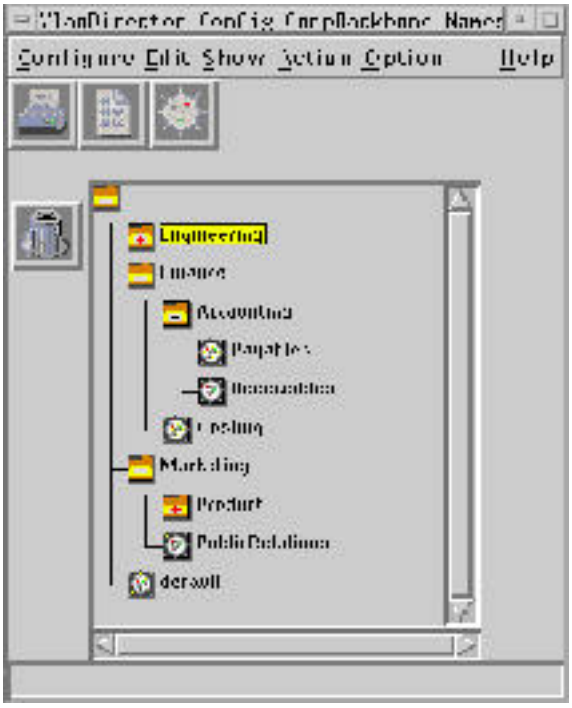


Figure 20. Topology Discovered by CDP

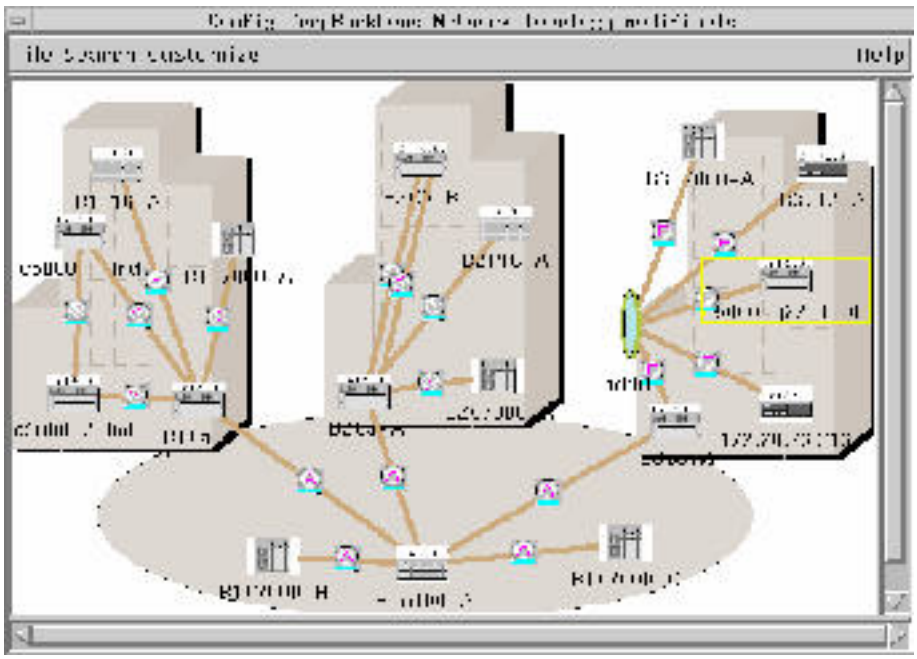
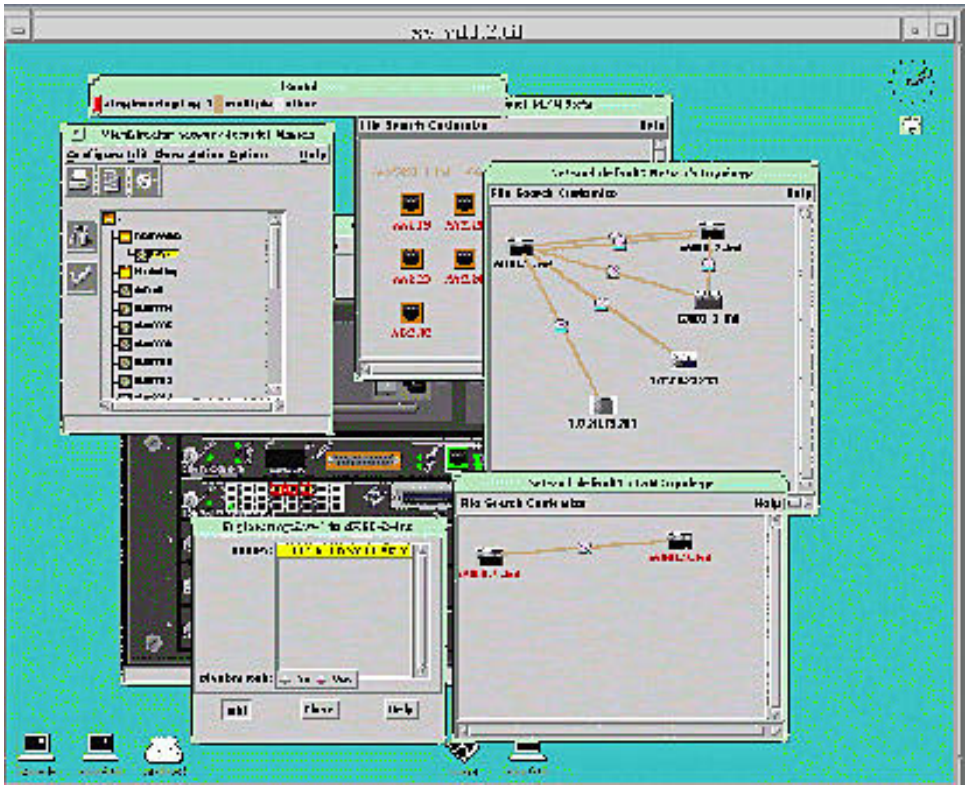


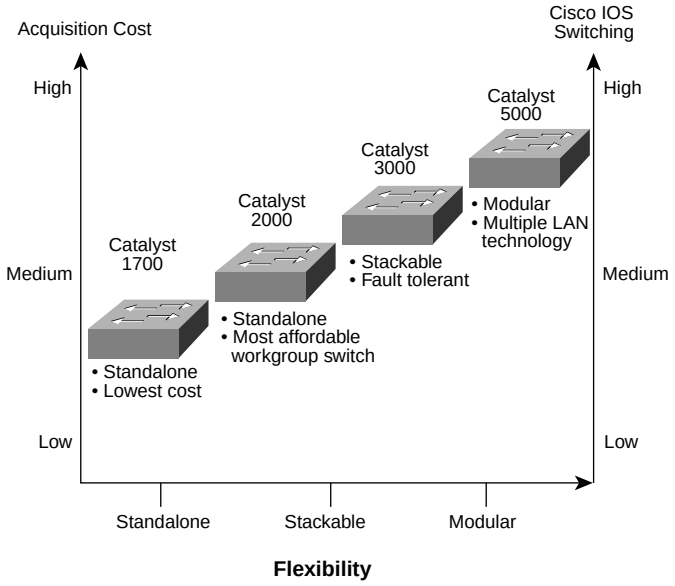
Figure 21. Manual Path Selection



Appendix D—Switch Architectures and Positioning

Three switching platform architectures are currently available: the Catalyst 1700/2000, the Catalyst 3000, and the Catalyst 5000. Although these switch architectures offer similar features and functionality, each contains products that are designed to serve a particular section of the market. Although a particular switch may solve a general networking issue, a lack of understanding of the full line of switches and their functionality could easily result in a switch being used inappropriately. Figure 22 summarizes the functionality and features of the three architectures.

Figure 22. Flexibility of Switch Architectures



The remainder of this appendix will describe each of the three switch architectures and their functionality.

Catalyst 1700 Desktop and Catalyst 2000 Family Workgroup Switches

The Catalyst 1700 is a standalone 25-port 10BaseT switch with two 100BaseTX ports. Designed as a high-performance alternative to a 10BaseT hub, the switch supports a single MAC address on 26 ports and unlimited addresses on a single backbone port. This backbone port can be either the 25th 10BaseT port or one of the 100BaseTX ports.

The Catalyst 2000 family includes the Catalyst 2100 and the Catalyst 2800 switches. The standalone Catalyst 2100 is equipped with 25 switched Ethernet ports, providing 10 Mbps of bandwidth to workstations or 10BaseT hubs. Two 100BaseTX ports provide a 100-Mbps link to servers or backbones. The switch supports 1024 MAC addresses per system with no per-port address limitations.

The standalone Catalyst 2800 also offers 25 switched Ethernet ports, but it comes with two high-speed expansion slots. Field-pluggable modules provide configuration, wiring, and backbone flexibility with a choice of 100BaseTX/FX and FDDI/Copper Distributed Data Interface (CDDI) modules available. An ATM module is also planned for connectivity to ATM backbones. ATM LANE will be supported with the release of the ATM module. Two versions of the product are available; one supports 2048 MAC addresses per system and one supports 8192 MAC addresses per system. There are no per-port address limitations on these switches.

The Catalyst 1700 and 2000 series is based on the Cisco-designed ClearChannel™ architecture, which offers wire-speed bridging, extremely low latencies, and nonblocking performance on all ports. This architecture employs a shared-memory scheme using three MB of DRAM. This design provides a large packet buffer as a shared-system resource for dynamic allocation of packet buffer memory to individual ports. All packets are received into and transmitted from memory. A packet remains in the same location in memory until all the ports forwarding that packet have completed their transmission. All address learning, address matching, packet queuing, packet buffer allocation, and forwarding decisions are performed entirely in hardware. The Forwarding Engine sits on a 48-bit-wide, 1-Gbps data bus, arbitrated by an out-of-band request scheduler. Transactions on the Packet Exchange Bus are scheduled by a combination of time-of-arrival, transaction priority, and port priority.

The ClearChannel architecture implements three distinct switching modes: FastForward™ (cut through), FragmentFree™ (modified cut through), and Store-and-Forward. The hardware implementation of the Forwarding Engine provides for high-performance, low-latency cut-through switching. It takes 31 microseconds (FIFO) to switch between 10BaseT ports and 7 microseconds (FIFO) between 100BaseT ports. The switched 100BaseT ports on the Catalyst 1700 and 2000 series can be configured for full duplex, providing up to 200-Mbps throughput.

The Catalyst 2000 series of switches support up to four port-based VLANs per switch. The port-based VLANs on these switches do not support ISL.

Catalyst 3000 Stackable Switching System

The Catalyst 3000 is a 16-port 10BaseT switch that has two open expansion bays. These slots can be populated with 100BaseTX/FX, 10BaseFL, 10BaseT, 100VG-AnyLAN, or ATM. The Catalyst 3000 also can be "racked and stacked." With the Matrix module, up to eight Catalyst 3000 switches can be stacked together as one logical switching system. A fully loaded Catalyst 3000 system can support up to 192 10BaseT ports or 128 10BaseT ports with 16 high-speed ports (any combination).

The Catalyst 3000 is based on the AXIS Bus architecture. The AXIS Bus is a 480-Mbps, time-slotted bus to which the LAN modules (or ports, both high-speed and 10-Mbps) connect. Each Catalyst 3000 or Catalyst 3000 System can support up to 10,000 addresses per box with 1700 addresses supported per port. Since address lookup is distributed to the ports, the master table is removed from the system as a potential bottleneck. Additionally, forwarding decisions are made at the port level, allowing high-performance, low-latency switching (< 40 microseconds, FIFO). In order to avoid packet loss in the event that the port of exit is busy, 384 Kb of memory is provided on each 10-Mbps port (192 K in each direction, input and output) and 512 Kb of memory is provided on the 100BaseTX/FX, 100VG, and ATM modules (256 K in each direction, input and output).

One of the primary features of the Catalyst 3000 System is the Matrix. The Matrix allows connection of up to eight Catalyst 3000 switches to each other to form one logical switching entity with one IP address. Interbox communication takes place via the StackPort module, a 280-Mbps port inserted into the back of the Catalyst 3000. This port also has an address table that makes forwarding decisions as to whether packets are forwarded on to the Matrix.

The Catalyst 3000 Matrix uses a simple crossbar switch fabric arrangement with round-robin output port arbitration. The crossbar switch matrix is a design by which each port is connected to every other port in the switch, allowing for low-latency switching. The switch and all links operate at the same clock rate, which is supplied by the Catalyst 3000 Matrix. All buffering is performed on the Catalyst 3000 StackPort interfaces. The Matrix is an unintelligent box; no processor is resident within. It is managed via the StackPort interfaces on the attached Catalyst 3000 switches.

The Catalyst 3000 supports up to 64 VLANs within the stack. This switch can validate 1024 VLANs in order to maintain compatibility with the Cisco 7x00 and Catalyst 5000. The Catalyst 3000 also supports ISL. ISL will run over 100BaseTX/FX and can be used to trunk multiple VLANs to another Catalyst 3000 System, Catalyst 5000, or Cisco 7x00. ISL will be supported in Q2 '96 and will be compatible with ISL implementation on the Catalyst 5000 and Cisco 7x00. ATM can also be used as a trunking protocol between switches. The first implementation of ATM for the 3000 system will support PVC with LANE support soon after. In the LANE release, LEC will be supported.

Catalyst 5000 Five-Slot Modular Switch

The Catalyst 5000 is a five-slot modular chassis with optional redundant power supplies. The first slot is dedicated to the Supervisor Module. This module is the switching engine and management board for the C5000 system; it also features two full-duplex 100BaseTX ports for file server connectivity or connectivity to other 100BaseTX switches. The four remaining slots can support:

- 24-port 10BaseT module
- 12-port 10BaseF module
- 12-port 100BaseTX module
- 12-port 100BaseFX module
- 2-port 100-Mbps FDDI/CDDI module
- Single-port 155-Mbps OC-3 ATM module

Figure 23. Catalyst 5000

The Catalyst 5000 architecture (Figure 23) is based on a high-speed single-bus switching fabric using an Output/Input Queuing Model. This cost-effective architecture gives the most efficient switching fabric for unicast applications today and multicast applications tomorrow. The Management Bus is a serial bus that carries configuration information to each module and statistical information from each module to the Network Management Processor (NMP). The Index Bus carries port-select information from the central Encoded Address Recognition Logic (EARL) to the ports. This information determines which ports forward the packet and which flush it from the buffer.

This architecture is strictly a store-and-forward operation; it uses central bus arbitration and address recognition logic for all modules. Each frame traversing the Switching Bus may be destined to a single port or to multiple ports, allowing for high-speed multicast forwarding without the need for frame copies.

The Switching Bus operates at 1.2 Gbits per second, a rate achieved by using a 48-bit-wide bus with a 25-MHz clock. The bus resides on the backplane with interfaces to each line module. Each port on each module has direct access to the bus through the 192-pin Future Bus connector on each slot. Through the Bus Arbiter, the bus supports a three-level priority request scheme. The bus also allows each port to perform a local flush and maintains a packet retry mechanism used during outbound port congestion.

Catalyst architecture has three basic components: the Bus Arbitration and EARL; the Port Interface; and the NMP. The Bus Arbitration and EARL, shared by all ports, together govern access to the data-switching bus and control packet transfer destination. Each Ethernet Port Interface comprises a custom ASIC, the SAINT, with an integrated 10/100 Ethernet MAC controller. Other media ports use a second custom ASIC, the SAGE, without an integrated MAC. To maximize current Cisco IOS software investments, the Catalyst 5000 uses a Motorola 68000 family processor.

Summary

Table 8 summarizes the configuration and performance characteristics of the Cisco Catalyst family.

Table 8. Summary of Cisco Catalyst Family

	Catalyst 1700	Catalyst 2100/2800	Catalyst 3000	Catalyst 5000
Configuration				
Max. Ethernet segments (w/o uplink)	25	25	192	96
Max. Ethernet segments (w/ uplink)	27	27	128 (w/ 16 uplinks)	96/72
Max. Fast Ethernet ports	2	2/16	16	50
Max. ATM ports	N/A	2 (2800)	16	4
Max FDDI ports	N/A	2 (2800)	N/A	4

Table 8. Summary of Cisco Catalyst Family (Continued)

	Catalyst 1700	Catalyst 2100/2800	Catalyst 3000	Catalyst 5000
Modular/Stackable/Standalone	Standalone	Standalone	Stackable	Modular
Performance				
Architecture (Cut through/Store-and-Forward)	CT, S/F	CT, S/F	CT, S/F	S/F
Latency	31 msec	31 msec	< 40 msec	34 msec
Backplane Capacity	1 Gbps	1 Gbps	4.8 Gbps	1.2 Gbps
Buffer Capacity	3MB	3MB	192 K per port	192 K per port
Features				
Addresses per Port	1	1024 (2100), 2048/8192 (2800)	1700	163
Addresses per Box	27	1024 (2100), 2048/8192 (2800)	10,000	16,000
Spanning-Tree Protocol (IEEE 802.1d)	No	Yes	Yes	Yes
Routing Functionality	No	No	No	No
IP Fragmentation	No	Yes (2800)	No	Yes
IP Multicast	No	No	No	Yes
ApaRT	No	Yes (2800)	No	Yes
Route Group (intergroup/intragroup)	No	No	No	Intergroup
VLANs	No	Yes	Yes	Yes
Access Lists	No	No	No	Yes
Address Filtering	Yes	Yes	Yes	Yes
Broadcast Suppression	No	No	No	Yes
SNMP	Yes	Yes	MIB II	MIB II
RMON	External	External	Yes	Yes
Ethernet Media (UTP[1], AUI[2], Fiber, BNC)	UTP, AUI, BNC	UTP, AUI	UTP/Fiber	UTP/Fiber
Topologies	E-net, FE	E-net, FE, FDDI	E-net, FE, ATM, 100VG	E-net, FE, ATM, FDDI
Redundant Power Supplies	No	No	Yes (w/ stack)	Yes
Hot Swappability	No	Yes (2800)	Yes (w/ stack)	Yes

1. Unshielded twisted-pair.
2. Attachment unit interface.

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).
0196R