The Architecture of Switched LANs

A TECHNOLOGY WHITE PAPER

AUTHOR: Martin Taylor, Madge Networks, March 1996

Executive Summary

This white paper is aimed at network planners and administrators who are concerned with planning the evolution of LAN installations to embrace the new generation of LAN switching technology. It is principally concerned with addressing the roles of switching and routing within the LAN, and the design of switched LANs around these two functions. The paper argues that switching should be treated as an opportunity for strategic upgrade of the LAN rather than as a tactical adjunct to routing. By connecting all LAN segments together with LAN switches and linking routers to the LAN through these switches, we can successfully optimize the performance/cost characteristics of the LAN, while at the same time improving flexibility and simplifying management.

About the Author

Martin Taylor is Vice President, Network Architecture for Madge Networks, Inc. He is responsible for the strategic planning of the company's products, including LAN switching hubs, ATM switches and LAN-ATM access switches. Martin joined Madge Networks in 1991, and formerly held the position of Director of Product Marketing. Prior to joining Madge, Martin held business development positions in Local Area Networks and in fiber optic cabling systems with GPT Ltd, the UK's leading supplier of telecommunications equipment. Martin has a total of 9 years experience in the communications industry, and a further 6 years experience in MIS management. He can be reached on email at mtaylor@madge.com.

Table of Contents

The Architecture of Switched LANs 7

Switching and Routing: Definitions 8

Network Design Goals 9

What is the Role of Routing in the Switched LAN? 10

Router-centric LAN Architectures 18

Virtual LANs 26

Controlling Broadcasts with LAN Switches 28

Routing: Where Does It Belong? 30

ATM in the Switched LAN 33

Conclusion 36

The Architecture of Switched LANs

Large LAN installations invariably consist of multiple LAN segments linked together by various kinds of internetworking devices – such as bridges, routers or switches. As constantly growing user demand for LAN bandwidth must be reconciled with the fixed 10 or 16 Mbps bandwidth of Ethernet or Token Ring segments, so the number of distinct LAN segments in a typical LAN installation is set to grow dramatically.

LAN switching has rapidly established itself as the technology of choice for boosting LAN capacity through increased LAN segmentation. This is because LAN switches can provide high performance transfer of LAN packets between and among large numbers of LAN segments, at low cost. Just as routers, in the past, have largely displaced bridges for interconnecting LAN segments, so switches are tending to displace routers in the LAN. But how far can this trend go? Is it possible to do without routers in the LAN?

There is no question about the need for routers to provide physical connectivity and protocol translation to interconnect switched LANs across the WAN. But opinions begin to diverge when we seek answers to questions such as:

- What is the role of routing within the switched LAN?
- What is the optimum balance of switching and routing in the switched LAN?
- What is the best way of physically interconnecting switches and routers in the LAN?

This paper sets out to provide rational answers to these questions.

Switching and Routing: Definitions

Before embarking on a discussion of the roles of switching and routing within the LAN, it is as well to be clear about the distinction between these two technologies.

LAN switches are like bridges. Generally they interconnect LAN segments of the same type, i.e. all Ethernet or all Token Ring. They may pass packets between ports transparently, or, if we are referring to Token Ring, by means of Source Routing. Transparent switches are not visible to end stations: they learn where stations are by monitoring all the packets on the LAN segments attached to their ports, and they direct packets to the appropriate ports according to the destination Ethernet or Token Ring address in each packet. This also means that they operate independently of the networking protocol that is being used by the end stations to communicate with each other – whether it be TCP/IP, Novell IPX, NetBIOS or IBM SNA. Source Routing switches for Token Ring differ from transparent switches only in that packets are directed to ports across the switch by information inserted into each packet by the end station, but again this is independent of the networking protocol.

In some cases, switches may be used to interconnect dissimilar LAN technologies. For example, some switches can interconnect Ethernet segments across an FDDI backbone. In this case, the switch is performing a simple translation of frame format from Ethernet to FDDI. This is done in such a way as to preserve the principle of transparency to the end stations.

Routers, on the other hand, are designed with the capability to relay packets from almost any type of network to any other. They are not transparent to end stations: indeed, an Ethernet end station that wishes to communicate with another end station the other side of the router actually addresses its Ethernet packets to the router, not the intended destination station. When a router receives a packet from one Ethernet LAN segment to be relayed to another, it strips the Ethernet header off the packet, examines the network address specified in the protocol header, and then looks up a table to determine whether this destination lies locally on one of its attached LAN segments, or whether the packet has to be sent on to another router. Having made this determination, the router will attach a new Ethernet header to the packet and send it on its way.

Routers maintain complex look-up tables to enable them to determine on which port to forward each packet. These tables are constructed by each router in cooperation with the other routers in the network, which pass information to each other about the state of paths through the network. The protocols and processes involved in this route determination are complex, and require a good deal of computing power and memory.

To summarize the essential difference between switching and routing within the LAN: packets passing through routers undergo a great deal more processing than packets passing through switches. As a result, routers tend to cost a lot more than switches for a given level of performance. In addition, each packet tends to spend a lot less time passing through a switch than through a router, and the lower latency of switches provides additional performance benefits. On the other hand, the processing power of routers can be harnessed to provide a greater degree of control than is usually possible with switches.

Network Design Goals

We cannot usefully discuss the optimum combination of switching and routing technologies for building switched LANs without identifying the design goals for the network. The following are some typical design goals for switched LANs:

- High aggregate capacity for reasonable cost
- Low end-to-end latency
- Flexibility to accommodate changing traffic patterns
- Ease of configuration and set-up
- Minimal administrative overhead associated with moves and changes
- Effective control of access to networked resources

In the discussion that follows, it will become clear that the majority of these goals are best met with a LAN design in which switching is the dominant technology, and where routing plays an important, but minor role. A high proportion of switching in the mix is generally desirable because switching offers greater traffic capacity at lower cost than routing, and is inherently easier to install, configure and manage.

What is the Role of Routing in the Switched LAN?

There are four principle functions which may be performed by routers within the switched LAN. A clear understanding of these functions provides useful insight into the role of routing within the switched LAN. The four functions are:

- Dividing the switched LAN up into broadcast domains, and linking these domains together
- Packet transfer between different subnets
- Interconnecting different LAN technologies
- Providing security of access for LAN-attached resources

These are certainly not the only functions that are performed by routers. When used to connect the LAN with the WAN, routers carry out a variety of protocol translations, for example to Point-to-Point Protocol for private line or dial-up connections, or to Frame Relay. They may also implement functions such as Data Link Switching to enable SNA traffic to be encapsulated in IP. But these functions are specific to WAN connectivity, and we are concerned here only with what happens in the switched LAN. Hence our focus on the four functions identified above.

Dividing the Switched LAN into Broadcast Domains

LAN technologies such as Ethernet and Token Ring provide the ability for any station to send a packet which will be received by all other stations on the LAN – a process known as broadcasting. Almost all networking protocols used on LANs make use of broadcasts to implement operational and administrative mechanisms, such as enabling client stations to locate servers, and allowing networked resources to disseminate information about available services.

In general, it can be said that the more stations there are attached to a LAN, the more broadcast traffic will be generated. This is true also when large LANs are constructed by linking many LAN segments together with bridges or switches.

Broadcast Protocol	Source	Purpose	Typical Frequency
Service Advertising Protocol	NetWare servers	Informs other NetWare servers about available services	Once per minute from each server
Service Advertising Protocol	NetWare clients	Locates nearest server	When client shell is loaded
Routing Information Protocol	IP and IPX routers	Informs other routers about network topology	Once per minute from each router
Routing Information Protocol	NetWare Clients	Learn NetWare network number	When client shell is loaded
Address Resolution Protocol	IP stations	Learn the MAC address associated with a known IP address	When an IP-based application is loaded
NetBIOS Add Name Query	NetBIOS stations	Ensure no duplicate name exists	When a NetBIOS-based application is loaded
NetBIOS Name Query	NetBIOS stations	Learn the MAC address	When a NetBIOS-based application is loaded

Figure 1: Common Broadcast Protocols within the LAN

Background Broadcast Traffic

The level of broadcast traffic that is observed within a LAN will depend not only on the number of stations that are attached, but also on a range of other factors such as the number of servers and routers on the LAN, the types of protocols in use, and the frequency with which users start up and shut down networked applications. Also, the observable broadcast characteristics of Token Ring LANs differ from Ethernet, in that Token Ring has the concept of source route explorer frames which may get replicated when there are multiple choices of route through a bridged network.

Because of the variety of factors that affect background broadcast traffic levels in the LAN, it is difficult to give clear general guidelines. However, practical measurements have shown that, even with flat bridged or switched LANs containing many hundreds or even thousands of nodes, the average level of background broadcast traffic is typically no higher than 10-30 packets per second, with occasional bursts of up to 100-150 packets per second. Since 30 broadcast packets per second represents around one quarter of one percent of Ethernet wire speed (based on average broadcast packet size of 100 bytes), the impact of this broadcast traffic is on network performance is negligible.

While these levels of broadcast traffic have negligible impact within the LAN, the same cannot be said about WAN connections. Background broadcast traffic can absorb a significant proportion of costly WAN bandwidth on low-speed leased lines, and routers perform a valuable role in minimizing the impact of broadcasts in this context.

Current trends in the type and usage of networking protocols and software are tending to reduce the amount of background broadcast traffic seen in the LAN. For example, the usage of NetBIOS, a particularly broadcast-intensive protocol, is declining. Also, new features incorporated by Novell in NetWare 4.x, including NetWare Directory Services and support for the NetWare Link State Protocol, dramatically reduce the amount of Service Advertising Protocol (SAP) and Routing Information Protocol (RIP) traffic that is seen in NetWare installations.

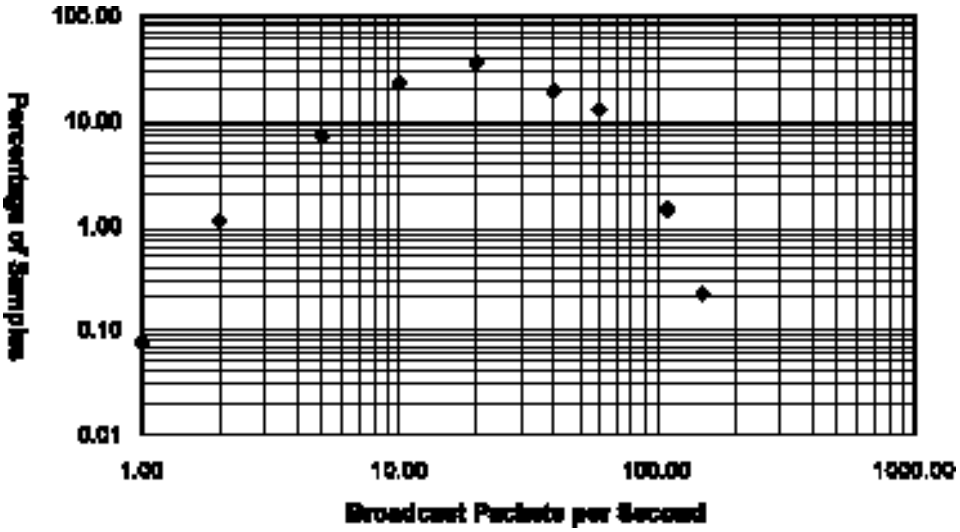

Figure 2: Frequency of Broadcast Packets Measured on a 3,000 Node Flat Network (Principal Protocols: NetBIOS, IPX)

Broadcast Storms

Long-serving network administrators may remember a type of problem that arose in the early days of widespread LAN deployment known a “broadcast storm”. This is a kind of chain reaction which could occur in large networks where a high level of broadcast traffic could temporarily saturate parts of the network, causing stations to lose connections to servers, triggering more broadcast traffic as these stations sought to re-establish their connections. The resulting rapid increase in broadcasts could cause growing saturation throughout the network, precipitating a catastrophic loss of communication.

The introduction of routers is widely credited with providing the solution to the problem of broadcast storms. Broadcasts which are transmitted by clients to find servers are not propagated across routers, because a client that needs to communicate with a server on the other side of a router actually sends the packet to the router for onward transmission. Thus the router provides a kind of “broadcast firewall”, which limits the propagation of broadcasts, and prevents the chain reaction which might trigger a broadcast storm.

Fear of broadcast storms has often led to a very router-centric approach to LAN design – a trend which the influential router vendors have done much to encourage. We illustrate what is meant by a router-centric architecture later in this paper.

There is no question that the broadcast storms experienced in large LANs in the days of bridged LAN architectures could cause serious loss of network service. However, the problems arose primarily from three factors which are largely absent today:

The use of remote bridges to link external sites over low speed leased lines. The original remote LAN bridges incorporated little or no broadcast filtering capabilities, and levels of broadcast traffic that used a negligible proportion of Ethernet’s 10 Mbps bandwidth could quickly saturate a 64 kbps line. The resulting loss of station connectivity could easily trigger a broadcast storm. Modern practice uses routers to support low speed connections to remote sites, and these routers prevent the remote links from becoming saturated with broadcasts.

Idiosyncrasies in the implementation of IP protocol stacks in end stations. The literature on IP identifies a number of historical implementation issues with end station protocol stacks that could cause broadcast storms. Among these are the requirement in some earlier versions of Berkeley Unix that stations should forward packets sent to them with an incorrect IP address, and the possibility that end stations could send Internet Control Message Protocol (ICMP) error messages in response to certain kinds of broadcast packets. Current versions of IP implementations have eliminated these issues.

Poorly performing implementation of the network interfaces and protocol stacks in end stations. Historically, a combination of insufficient processing power, inadequate buffer memory and immature software implementations of protocol stacks resulted in excessive sensitivity to broadcast traffic at LAN stations. With LAN interfaces becoming congested at relatively low levels of broadcast traffic, connections could be lost, and the resulting attempts by stations to re-connect could create the conditions for a broadcast storm. With the benefit of ten years maturing of the technology, LAN interfaces can now handle very high rates of broadcasts without becoming congested. The threshold level of broadcast traffic that might risk triggering a storm has thus been raised by orders of magnitude.

In summary, the risks of broadcast storms within switched LANs today are often greatly exaggerated. With a modest amount of attention paid to the correct configuration of LAN switches, there is no reason why large “flat” switched networks containing thousands of nodes cannot be constructed – with very substantial benefits in terms of price/performance and flexibility. More on this later.

Packet Transfer Between Subnets

Widely used networking protocols such as IP, IPX and NetBIOS provide a Network Layer addressing structure which is largely independent of the underlying LAN transport. Both IP and IPX are known as routable protocols. This means that they implement a hierarchical addressing scheme of the form <network id • host id> to identify all networked hosts. NetBIOS is known as a non-routable protocol, since networked hosts are identified simply by a name which has no hierarchical structure.

The addressing structure of networking protocols has important implications for the design of switched LANs, because the hierarchical nature of the addresses requires that networked stations be divided into groups that have the same <network id>, and the only way that a station in one group can communicate with a station in another group is to send packets to a router for forwarding. This section therefore describes addressing schemes in considerable detail, while another section later in this document describes strategies for working around some of the limitations of these schemes.

IP Addressing

IP uses a total of four bytes (32 bits) for Network Layer addressing. The split of the number of bits between <network id> and <host id> is somewhat flexible. Organizations that use IP can choose to employ a private addressing scheme, in which case they have a great deal of flexibility in the way that the network addresses are structured, or they can use the public scheme administered by the Internet Assigned Numbers Authority (IANA) which allocates blocks of globally unique IP addresses.

Most organizations using IP on a large scale choose to use a public addressing scheme. The problem here is that with only four bytes, the address space for creating globally unique addresses is very constrained. As a result, many organizations are forced to use an address format that places serious limitations on the number of stations that can intercommunicate using IP on a LAN without having to pass via a router.

With routable protocols, each end station has a network address consisting of <network id • host id>. For IP, the address of each end station is normally manually configured by the network administrator. When an end station wishes to communicate with another end station whose IP address it knows, it compares the <network id> of the destination station with its own <network id>. If they are the same, this means the destination station is effectively on the same LAN, and we just need to find out its LAN address, for which we use the Address Resolution Protocol. If they are not the same, then the stations will have to communicate via one or more routers. Routers contain information about how to reach all the different network ids.

What this all means in practice is that routers may be needed in the switched LAN to enable end stations with different network ids in their IP addresses to intercommunicate. With the commonest scheme in use, Class C addressing, we have to divide LAN users into groups of not more than 254 stations with the same network id. Communications within these groups may take place directly across the switched LAN, while communications between the groups must take place via a router.

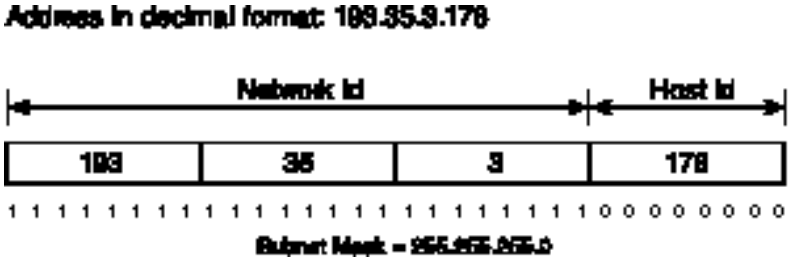

Figure 3: Composition of Typical Class C IP Address

IP Addressing: Subnetting

While the most commonly used IP addressing scheme, Class C, suffers the limitation that no more than 254 stations can belong to the same subnet, other IP addressing schemes are not so limited.

Many larger organizations are fortunate enough to have been allocated one or more Class B or even Class A addresses by the IANA. A Class B address configured as a single subnet supports up to 65,534 stations, while a Class A address allows over 16 million stations on a single subnet. The IANA is now extremely reluctant to allocate new Class A or B addresses because there are so few remaining unallocated, and a strong justification must be made in support of any application to the IANA for Class A or B addresses.

In practice, no real network is going to need as many as 65,000 stations in a single subnet, let alone 16 million. To make better use of the allocated address space, it is usual to configure a “subnet mask” that allocates some of the <host id> address space to the <network id>. With a Class A address, the <network id> portion of the IP address is normally the first 8 bits of the IP address, while for a Class B address, the <network id> is the first 16 bits of the IP address. But if we allocate an additional 6 bits, for example, from the <host id> space to a Class B address, then we effectively split the Class B address into 62 smaller subnets with up to 1,022 stations per subnet. (We might expect to get 64 subnets out of this Class B address by adding 6 bits to the <network id>, although in fact we only get 62, for reasons that are too complex to explain here).

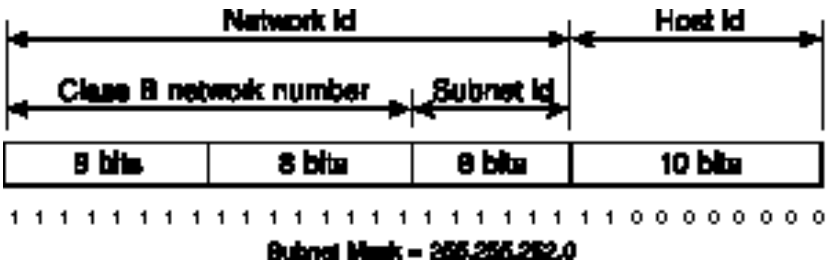

Figure 4: Composition of a Class B Address Divided into 62 Subnets

Organizations that own Class A or Class B IP addresses have a valuable asset that should be used wisely. Setting the size of the subnet mask is critically important to an organization's ability to make the best use of its addressing space. The subnet mask should be large enough to support the maximum number of distinct subnets required (for example, there may need to be one subnet per branch office), while at the same time allowing for large enough subnets in major sites to enjoy the full price/ performance benefits of switched LANs.

The worldwide pool of Class A and Class B addresses is extremely limited. Many organizations applying to the IANA now are likely to have to work with Class C addresses if they want access to the Internet. Strategies for working around the limitations of Class C addressing are discussed later.

IPX Addressing

Novell's IPX protocol suffers few of the addressing limitations seen with IP. In the IPX protocol, the Network Layer address occupies ten bytes, of which the first four bytes is the network id and the last six bytes is the host id. The host id is in fact simply a copy of the burnt-in address of the LAN adapter, which itself is a globally unique number allocated by the Institute of Electrical and Electronic Engineers (IEEE). The network id is obtained by each station when it is started up by broadcasting a request to the NetWare servers on the LAN.

Therefore with IPX, there is effectively no limitation on the number of stations which can have the same network id. Within a correctly designed switched LAN, all IPX stations can freely intercommunicate across the LAN switches without having to go through a router.

NetBIOS Addressing

NetBIOS stations are addressed simply by an alphanumeric name which has no inherent hierarchical significance. Therefore NetBIOS stations need a flat or bridged LAN across which to intercommunicate. If forced to go through a router, NetBIOS is either bridged by the router, or encapsulated in another protocol such as IP.

Network Reality – The Multiprotocol LAN

Most real LANs in large organizations are required to support a mixture of LAN protocols. However, the differing characteristics of each protocol's addressing scheme suggests a different optimum LAN design for each protocol. Fortunately, it is possible to design a switched LAN that provides the optimum performance for IP, IPX and the non-routable protocols.

The traditional approach to LAN design has been IP-centric, and has focused on the necessity to divide the LAN into subnets as dictated by the IP addressing scheme. Many organizations have taken the view that, if they have to work with a maximum of 254 stations per subnet, they may as well design the LAN around one subnet per physical LAN segment. This leads to the classical collapsed backbone router architecture, where every LAN segment is connected to a router port.

The problem with this approach is that it requires all traffic between LAN segments, regardless of protocol type, to pass through one or more routers. In effect, we have taken a network structure suggested by the limitations of the IP addressing scheme, and imposed this structure on all the other protocols. This ignores the possibilities offered by IPX to allow much larger numbers of stations within each subnet, all intercommunicating without going through a router. It also implies that all non-routable protocols such as NetBIOS must be bridged through the routers. The router-based collapsed backbone therefore requires a large amount of costly router capacity to achieve adequate performance.

In summary, the necessity for routers to transfer packets between multiple subnets in the switched LAN is imposed only by the limitations of the IP addressing scheme, and does not apply to IPX and to the non-routable protocols. With the correct design of switched LAN, the amount of routing capacity within the switched LAN need only be sufficient to handle the IP traffic passing between stations with different network or subnet numbers. In most cases it is possible to pass all local IPX traffic and all non-routable traffic everywhere within the switched LAN without passing through a router. This will be explained in more detail in the section on switched LAN design below.

Interconnecting Different LAN Technologies

Routers are extensively used in large LAN installations to provide connectivity between different LAN technologies. Connections between Ethernet or Token Ring workgroups and an FDDI backbone, and links between Ethernet and Token Ring LANs on the same site are two common examples of this.

The role of the router in supporting connections to FDDI backbones is being challenged by LAN switches with translational bridging capabilities between Ethernet or Token Ring and FDDI. By focusing on simple frame format translation and eliminating all the complexities of Network Layer protocol processing, these devices offer FDDI backbone connectivity at a fraction of the cost of routers. And as FDDI gives way to ATM as the high-speed backbone technology of choice, the task of connecting Ethernet or Token Ring workgroups to the backbone with switches becomes even easier – since ATM with LAN Emulation supports the transport of Ethernet and Token Ring frame formats directly, without the need for translation.

Translation bridges for connecting Ethernet to Token Ring have a long history of interoperability problems, and routers remain the best choice for this need. Where there is a major need for Ethernet and Token Ring users to share access to common resources, however, the installation of both types of LAN adapter into each server, giving direct access to both classes of users, will provide far better performance and greatly reduce the need for costly router capacity.

Providing Security of Access

In addition to forwarding packets between different kinds of LAN and WAN connections, routers generally provide a range of packet filtering capabilities intended to provide secured access to networked resources. The necessity for secure access for WAN connections is obvious, but many organizations also take advantage of packet filtering in routers to secure access within the LAN. Routers can offer a useful supplement to the security functions that are normally provided by network applications, and can "hide" networked resources from users who do not have the authority to access them.

Most routers offer a range of logical rules that can be applied to create appropriate packet filters, based on network addresses, socket numbers, protocol types and so on. These rules provide great flexibility to enable users to set up precisely the security functions they require. However, it should be noted that filtering rules are processed in software for each packet received by the router, and this can have a serious impact on both router throughput and per-packet latency.

Router-centric LAN Architectures

In general, the network architecture strategies promoted by the leading router vendors have evolved through three clear historical phases:

Phase 1 – Elimination of all bridges and replacement with routers, beginning with the remote bridges and progressing to local bridges. This leads to a LAN architecture based on a collapsed backbone comprising one or more large multi-port routers, typically linked by a high speed LAN such as FDDI.

Phase 2 – With the supremacy of the collapsed backbone router threatened by a new breed of internetworking device, the LAN switch, the router vendors are forced to embrace LAN switching but promote the use of this technology as a workgroup solution feeding into collapsed backbone routers.

Phase 3 – In a bid to offer greater flexibility to network users, the router vendors promote the concept of “Virtual LANs” which allow the switched network to be divided up around the routers logically rather than physically. (The subject of Virtual LANs is discussed later).

The need to subdivide large LANs into many smaller broadcast domains, whether physically or logically, has been a central theme of the strategies promoted by router vendors. This should come as no surprise, since it requires routers to provide communications between the broadcast domains.

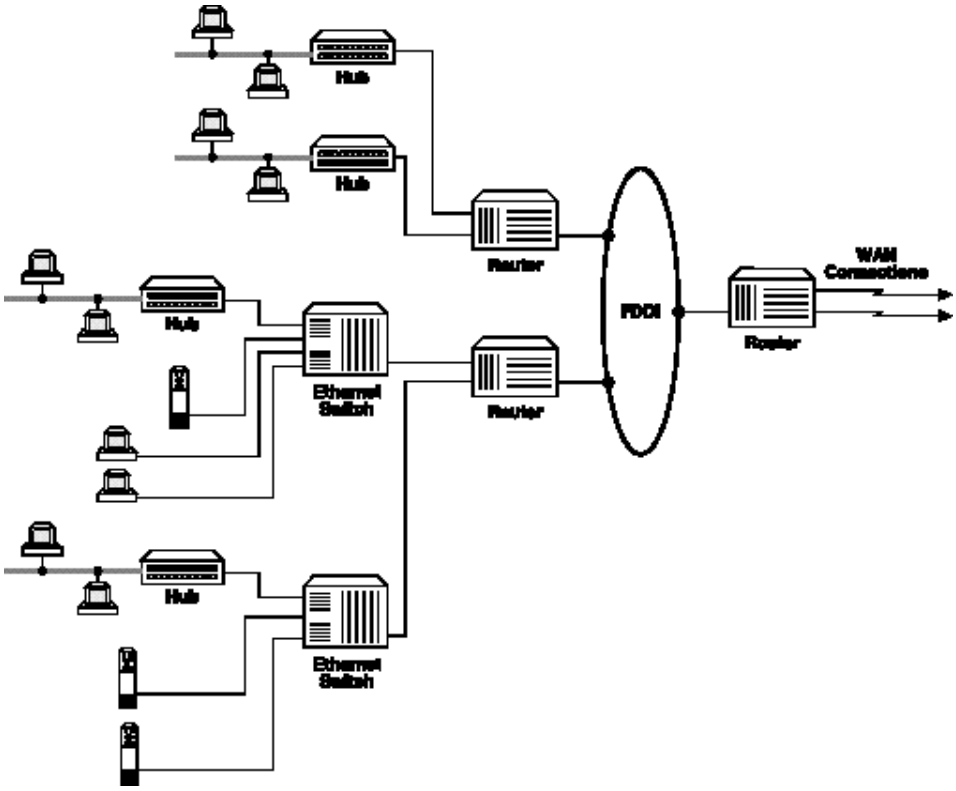

Figure 5: Router-centric Architecture for Switched LANs

There has never been any question about the superiority of routers over remote bridges for LAN-to-LAN connectivity over the WAN. The abilities of routers to prevent broadcasts from saturating low-speed WAN links, to dynamically re-route around failed WAN links, and to support protocol translation between LAN protocols and WAN protocols such as Frame Relay and Point-to-Point protocol, are all essential for a healthy and efficient network.

The historical penetration of routers into the LAN can probably be attributed to three key factors: fear of broadcast problems within the LAN; the growing importance of IP as a networking protocol; and the attraction of large, multi-ported devices for linking LAN segments, as compared with two-port local bridges.

While we need to be conscious of each of these factors, our judgment about the relative importance of routing in the LAN should be tempered by the new realities:

- Broadcast traffic within the LAN is less of an issue than it once was, and is in any case subject to other kinds of control besides the use of routers.
- While IP continues to become more prominent within the LAN, other protocols remain very important, and an IP-centric approach to LAN architecture may seriously disadvantage these other protocols.
- LAN switching is a mature and widely available technology that provides a more cost-effective means of interconnecting multiple LAN segments than large routers.

In the next section we shall discuss how these new realities guide us to an optimum architecture for switched LANs.

The Design of Switched LANs

The new approach to LAN design starts with the general principle that LAN switching offers higher traffic capacity at lower cost than routing. In order to maximize the cost-effectiveness of the LAN, this implies that we should aim to switch as much as possible of the traffic between LAN segments, and route as little as possible. And if we design the network according to this general principle, it turns out that we obtain a number of secondary benefits – as we shall see.

The number of stations per physical segment is, of course, an important parameter in the network design. There are no hard and fast rules about this – it may depend on a number of factors, including the type of LAN technology (Ethernet, Token Ring or other), the bandwidth needs of the networked applications, and the physical layout of the buildings. Local decisions must be taken based on experience, and judgment as to what constitutes acceptable application performance. The rest of the network design exercise is now concerned with how we link all the LAN segments together.

We can successfully maximize the amount of traffic that is switched, rather than routed, between LAN segments simply by connecting all LAN segments together via LAN switches. This means that there is always a switched path between any two points in the network, and it implies that we are not physically separating parts of the LAN from one another with routers.

With all LAN segments connected via switches, we have effectively created a single large broadcast domain in the LAN. As a result of this, we can expect to see significant levels of broadcast traffic within the LAN, although this will be firewalled from the WAN by the routers that are used to support the WAN connections. The overall level of broadcast traffic within the switched LAN is not likely to consume a significant proportion of the LAN bandwidth, even in LANs with thousands of nodes – although this will depend to some extent on the types of LAN protocol in use. In extreme cases, it may be necessary to take some simple measures to reduce the level of broadcast traffic within the switched LAN. These measures may include the configuration of Virtual LANs to break the switched LAN up into a number of broadcast domains, and the implementation of broadcast filtering in the LAN switches. We will come back to this subject later.

The Flat, Switched Network

A single, large broadcast domain within the switched LAN has the following implications for the common LAN protocols:

- With IPX, a broadcast domain is always a single network id or subnet. All servers within a broadcast domain must be configured with the same network id, and the clients obtain their network id information from the servers. Since all IPX stations have the same network id, they can all intercommunicate directly across the LAN switches, without passing through any routers.
- With IP, the <network id> portion of the network address in each station is usually manually configured, and stations that have different network ids in their addresses must communicate via a router. However, stations with different subnet numbers can co-exist within the same broadcast domain. This means that if the switched LAN comprises one single large broadcast domain, then stations with the same network id can be located anywhere within the switched LAN, and will intercommunicate without passing through a router. Stations with different network ids can only intercommunicate via a router, and therefore there must be one or more routers connected to the switched LAN to provide for this.
- With non-routable protocols such as NetBIOS, communication between stations never requires a router, and all stations intercommunicate directly via the LAN switches.

One obvious question that arises is: if we need a router to allow stations on different IP subnets to intercommunicate, and we have connected all the LAN segments together with switches, where do we connect the router? The answer is that the router or routers may simply be connected to one or more LAN switches by one or more ports at some suitable location in the switched LAN. Traffic that is to be passed from one station to another via a router is not constrained to pass between two different physical ports of a router – it can enter and leave the router via the same port, provided that the router port is configured with the details of all the subnets that are present in the switched LAN connected to this port. Depending on the level of traffic between IP subnets, the required amount of routing capacity may be provided by existing routers used to support WAN connections, or by separate dedicated “one-legged” routers connected to the switched LAN.

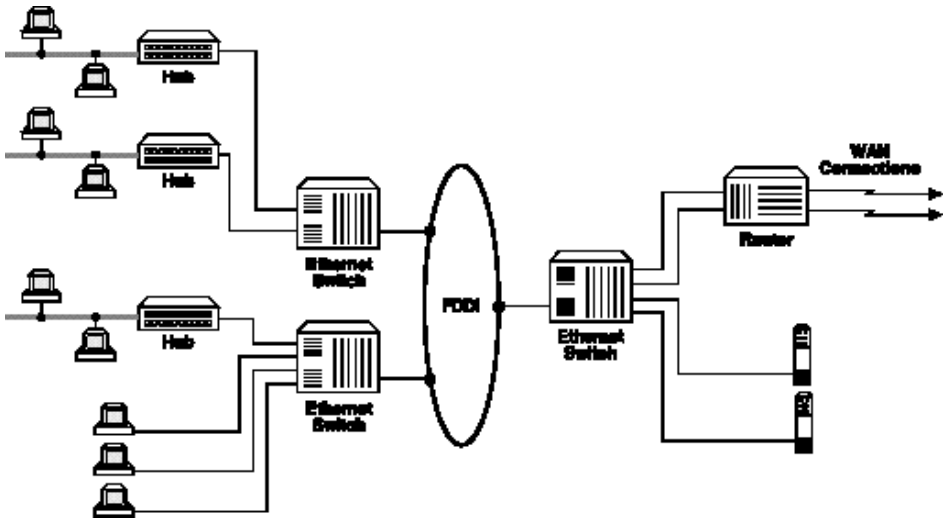

Figure 6: Router attached to a Switched LAN, for both IntraLAN and InterLAN Routing

Meeting the Network Design Goals

We have described a switched LAN in which all LAN segments are interconnected by LAN switches, and which comprises a single flat network. We have also seen how this meets the first goal, to provide high traffic capacity at low cost, by ensuring that the vast majority of the local traffic is switched rather than routed between LAN segments. But how well does this approach meet our other network design goals?

1. High aggregate capacity for reasonable cost
2. Low end-to-end latency
3. Flexibility to accommodate changing traffic patterns
4. Ease of configuration and set-up
5. Minimal administrative overhead associated with moves and changes
6. Effective control of access to networked resources

Low end-to-end latency is helpful in providing high throughput for data transfers, especially with latency sensitive protocols such as NetWare Core Protocol running over IPX. Low latency is also a requirement when we want to run networked multimedia applications. Provided they have been well designed, switches show considerably lower latencies than routers, with cut-through switches offering the lowest latency of all. Allowing all segments in the network to intercommunicate via LAN switches helps to minimize end-to-end latency in the LAN.

Flexibility to accommodate changing traffic patterns. Traffic patterns within the LAN are subject to constant change as new applications are deployed and user populations for each server move around in response to organizational changes. From the point of view of the network user, the performance he sees will vary according to the number of hops between the local LAN segment and the server being accessed. One or more routers between the user and the server can seriously degrade the observed throughput compared with access to a local server. A flat network structure based on linked LAN switches helps to reduce the number of hops between users and servers, and the performance characteristics of switches helps to deliver consistent performance wherever the users and servers are located in relation to one another.

Ease of configuration and set-up. Switched networks are easier to install, configure and manage because switches are inherently much simpler devices than routers. Routers require complex configuration and set-up to provide the desired connectivity for each protocol, and also require considerable tuning of parameters to obtain optimum performance. Switches, on the other hand, operate in essentially the same way for all protocols, and there is little to configure to get the network up and running. Therefore a LAN which relies primarily on switching to interconnect LAN segments reduces the effort involved in set-up and configuration.

Minimal administrative overhead associated with moves and changes. LANs in which all the segments are interconnected by switches reduce the effort involved in administering moves and changes. In the classical collapsed backbone router design, each physical segment is associated with a single IP network id. All of the stations connected to each physical segment must be configured with an IP address containing the matching network id. If a station is moved from one physical LAN segment to another, it is necessary to re-configure the station with a new IP address to reflect the change in network id. If, on the other hand, all the LAN segments belong to a single broadcast domain, then any station with any network id can be connected to any physical segment. It will always be able to communicate with other stations that share the same network id via the LAN switches, and with other stations that have different network ids via a router. With this design, therefore, there is no need to assign a new IP address whenever a station is moved. The initial assignment of IP addresses to stations is based not upon the physical location where the station is to be connected, but on the logical association of the user with the resources to which most frequent access is required.

Effective control of access to networked resources. The final design goal, to provide effective control of access to networked resources, is not specifically served by LAN switches. Indeed, connecting all LAN segments together with LAN switches provides for free access to all network resources to all users, at least in terms of network connectivity. In practice, for many applications, this is what is required. Access control is applied at the application level by passwords and user privileges that are set up by the system administrator. Where additional security is required, it becomes necessary to separate secured resources from the rest of the network by means of a router. This separation may be applied physically, where the the secured resources are attached to LAN segments connected to the rest of the network via a router, or logically, by defining a virtual LAN around the secured resources. Virtual LANs are discussed below.

Solutions to IP Addressing Limitations

Many organizations have been using IP in their LANs for a number of years, using an addressing scheme that was defined long before LAN switches and their implications were widely understood. It is not surprising, therefore, to find that the addressing scheme currently in use imposes undesirable restrictions on an organization's ability to maximize the benefit obtainable from LAN switching. This may be because the organization has been able to obtain only Class C addresses from the IANA, or because the organization originally chose a subnetting scheme to suit a router-centric LAN design, where each physical LAN segment is allocated its own subnet id.

Given the practical difficulties of day-to-day IP address administration, there is often considerable reluctance even to discuss the possibility of changing an existing IP addressing scheme. However, there are a number of options, which can not only bring improvements through better use of LAN switching, but which may also simplify the task of address administration itself.

Assign Subnets to Logical User Groups

When IP addresses are assigned to individual end stations, we have to decide to which subnet the station is to belong. In a router-centric LAN, it is common to associate a network id with an individual LAN segment or a physical area of the LAN site, and therefore to assign users to subnets on the basis of where they are located or to which LAN segment they are attached. This has two important implications. First, if the servers to which these users normally connect are located in a different part of the site or are attached to a different LAN segment, then the users will only be able to communicate with their servers across a router - and suffer the inevitable degradation in performance. And secondly, usage of the available IP addresses in the subnet address space may be very inefficient. For example, if we have 60 users on a LAN segment which is allocated its own Class C network id, then we are wasting 75% of the available addresses on this subnet.

We have described a flat, switched network in which any network id can be present on any LAN segment. This allows us to decouple the assignment of network ids to end stations from the physical location of the station within the LAN, and to focus instead on the logical communications needs of particular stations. By assigning groups of end stations and the servers to whom they commonly connect to the same network id, regardless of their physical location, we can ensure that the majority of users enjoy the performance benefits of switched connections to their "home" servers. Furthermore, we can make much better use of the available address space within any given network id.

In practice, this means that the assignment of network ids to end stations should be based on membership of a logical community rather than physical location within the LAN. For example, we might define all members of a particular department as belonging to a subnet. One useful result of this approach is that we don't need to assign a new address to an end station that is moved within the LAN. Provided the user still belongs to the same logical community, he or she will still benefit from switched connections to the servers within this community.

Multiple IP Stacks in Servers

By assigning end stations to subnets on the basis of logical communities of communication, we can substantially increase the proportion of LAN traffic that is carried on switched connections rather than having to be relayed by routers, with obvious benefits in terms of overall performance and capacity of the LAN. But there will be some cases where users from more than one IP subnet need frequent access to a particular server. This may happen if a range of different applications is being run on a very large server, or where large groups of users need to access common resources such as email servers.

An effective solution to this problem is to provide a direct connection between these servers and multiple subnets. This implies assigning multiple IP addresses to the server, so that it has a local presence on each of the subnets that it serves. Some operating systems, such as Microsoft Windows NT, allow multiple protocol stacks to be installed over a single network adapter card, so that one physical connection to the switched LAN supports multiple logical network layer connections. Others may require that a separate physical network interface be installed for each IP address.

CIDR - Classless Inter-Domain Routing

Classless Inter-Domain Routing (CIDR) is an enhancement to the basic operation of IP routed networks that sets out to deal with the problem of Internet address space shortages. CIDR emerged from the observation that Class B addresses offer an unnecessarily large number of possible station ids (65,534) for most organizations, while Class C addresses offer too few (254). There are large numbers of organizations for whom the optimum addressing scheme lies somewhere between these two extremes.

This led to the idea that such organizations could be assigned contiguous blocks of Class C addresses which share the same most significant bits in the <network id> portion of the address. So for example an organization which required a total of up to 4096 station addresses would be assigned a contiguous block of 16 Class C network ids, all of which would share the same bit values in the most significant 4 bits of the third octet of the address.

Just as Class B addresses can be divided into subnets by defining a subnet mask that is longer than the 16 bits of the base Class B address, so conversely blocks of contiguous Class C addresses can be “supernetted” or aggregated into larger subnets by defining a subnet mask that is shorter than the 24 bits of the base Class C address. In the example given, the 16 Class C network ids could be aggregated into a single subnet by defining a subnet mask with the most significant 20 bits set, or into 4 subnets of 1022 stations each by defining a 22-bit subnet mask.

The value of CIDR is that it allows organizations that have not been able to obtain a Class B address from the IANA the freedom to define subnets with more than 254 stations, and therefore to benefit more easily from the performance gains obtainable from switching in large LANs.

The routing protocols necessary to support the use of CIDR have been widely deployed since the spring of 1994. Nevertheless, it would be wise to check the degree of support provided by existing routers for CIDR when contemplating the deployment of this scheme. End station protocol stacks may also need to be checked: older IP implementations would not support a subnet mask of less than 24 bits for Class C addresses, although recent implementations, such as that embedded in Microsoft Windows 95, are compatible with CIDR.

Private Addressing Schemes

Another solution that eliminates the restriction of 254 stations per subnet inherent to Class C addressing is to use a private addressing scheme.

A portion of the IP addressing space has been set aside by the IANA for private addressing schemes. In effect, this allows organizations to use a Class A addressing scheme, which provides a very high degree of freedom to support large numbers of subnets with large numbers of stations per subnet. The only limitation is that these addresses cannot be used by stations that need direct access to the Internet. Where direct Internet access is needed, this must be provided by means of an address translation gateway, which also acts as a security firewall.

The freedom provided by what is effectively Class A addressing solves many of the problems caused by limited address space. It is particularly helpful in large organizations with a mix of large sites and many small sites, where even Class B addressing may not provide the right combination of enough stations per subnet on large sites with enough subnets to support all the small sites.

However, private addressing schemes introduce considerable extra complexity to Internet access, so organizations that are contemplating the use of private addressing should proceed cautiously if they have any expectation of a growing need for Internet access.

Virtual LANs

The concept of Virtual LANs, or VLANs for short, has been much touted by LAN switch vendors as a new way of exploiting switching technology to bring greater flexibility to the network.

With a classical collapsed backbone router-based LAN, all LAN segments are connected to router ports and the LAN is physically divided up into separate broadcast domains by the router itself. This means that each LAN segment must be treated as a separate subnet, and all communications between segments must traverse the router. For performance reasons, it is highly desirable for users to be able to communicate with their most heavily utilized servers without having to go through a router. Therefore most servers need to be connected locally to the LAN segments where their user communities are connected.

The result is a rather inflexible network design that does not adapt well to change. When user populations move in response to organizational evolution, traffic across the router can become excessive, and performance suffers.

The idea of Virtual LANs is to allow broadcast domains within the LAN to be defined logically rather than physically. In a switched network where all LAN segments are interconnected by switches, and where routers are also connected into the switches, the network administrator can define logical groupings of segments or end stations to comprise broadcast domains. These domains may be created by means of filtering broadcast packets within the LAN switches in such a way as to confine any broadcast originating within a particular VLAN to just those segments that make up this VLAN. (Some switches provide additional security by confining also unicast packets within the VLAN). The routers are then used to provide communications between the VLANs. Because the broadcast domains are defined logically, via the network management console, rather than physically, by separating LAN segments with a router, the network should be far more flexible to accommodate the changing traffic patterns that are the inevitable result of organizational evolution.

While the VLAN concept sounds attractive in principle, it remains problematic to put into practice. From the user perspective, the extraordinary flexibility afforded by VLANs has one very important downside: the administrative burden of defining and maintaining the VLAN membership lists. And from the vendor perspective, there are no standards in place for the additional protocols required to implement VLANs across switched LANs that have more than one switch.

Some efforts are being made to simplify and automate the configuration of VLANs to reduce the administrative burden. One approach is for the LAN switches to identify the IP network ids of the stations connected to them, and to set the boundaries of the VLANs to match the boundaries of the IP subnets. This means that when a user moves to a different segment, the switches will reconfigure his VLAN to follow him. The disadvantage of this technique is that the VLANs which are defined to suit IP are then applied to all the other protocols, forcing a lot of IPX and non-routable protocol traffic to pass unnecessarily through a router.

Much of the emphasis placed on VLANs is based on two implicit assumptions, both of which are questionable:

Switched LANs have to be broken up into a large number of distinct broadcast domains. The levels of broadcast traffic in the majority of LANs allow for broadcast domains containing many hundreds or even thousands of users to be configured without problems. Therefore, even in the largest installations, the number of distinct VLANs needed is likely to be small.

VLANs are necessary to support IP subnets within a switched LAN. As we have discussed earlier, there is no problem in allowing multiple IP subnets within a broadcast domain, and therefore there is no necessity to map VLANs directly to IP subnets.

When developing a policy for implementing VLANs, network administrators should carefully weigh up the potential benefits of VLAN deployment against the cost of administering and controlling the membership of each VLAN. As a general rule, VLANs should be made as large as possible consistent with maintaining an acceptable level of broadcast traffic. This will not only make it easier to decide which users and which servers should belong to each VLAN, it will also minimize the amount of routing capacity that needs to be installed in the LAN to support packet transfer between VLANs.

In addition to controlling the level of broadcast traffic within the switched LAN, VLANs can also be used to firewall off parts of the LAN for security purposes. Routers can then provide a secure link from these VLANs to the rest of the network.

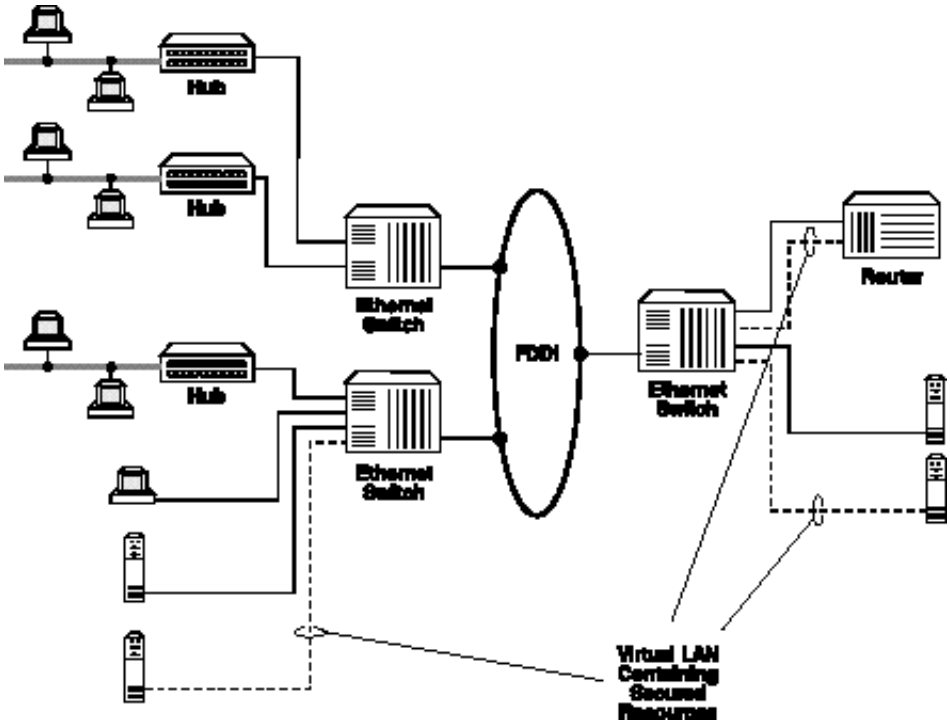

Figure 7: Secured resources on a Virtual LAN, linked via a router

Controlling Broadcasts with LAN Switches

Many LAN switches contain features that may be deployed to help control the amount of broadcast traffic in the LAN, enabling larger broadcast domains to be created and therefore maximizing the advantages of switching in the LAN. There are two kinds of features that are useful here: overlapping VLANs, and intelligent broadcast filtering.

Overlapping VLANs

The conventional concept of VLANs is that each VLAN represents a separate and exclusive broadcast domain. This means that each station can belong to only one VLAN. With this definition, the only way to pass packets between one VLAN and another is to send them via a router.

However, some LAN switches allow the configuration of overlapping VLANs. There are two variants of overlapping VLANs. "Fully flexible" overlapping VLANs allow each station to belong to one or several VLANs. Broadcasts sent out by a station will be visible to all the members of all the VLANs to which the sending station belongs. "Nested" overlapping VLANs provide for the definition of super VLANs that consist of groups of smaller VLANs. Stations defined as belonging to a super VLAN are able to exchange broadcast traffic with all the stations that are members of the sub VLANs, but broadcasts cannot propagate from one sub VLAN to another.

Both of these approaches allow for central servers to send to, and receive broadcasts from, multiple workgroups, while preventing the passage of broadcasts between one workgroup and another.

The use of overlapping VLANs in this way can substantially reduce the average amount of broadcast traffic that is seen on the individual LAN segments and at the end stations. At the same time, the common resources to which many users need access can be reached directly via switched connections, without having to go through a router. Not only that, but overlapping VLANs can provide a high degree of security, allowing users on many workgroups to access common resources while denying access to local resources for users from other workgroups. Thus we have achieved two of the main aims of router deployment – broadcast control and access security – simply by the intelligent use of switch features.

Overlapping VLANs also help with the connection of routers to the switched LAN. When we connect a router to a LAN switch to provide a link between VLANs, we might expect to have to connect as many physical router ports as there are VLANs to be linked. This would be the case if the VLANs are the conventional, exclusive type. If we can define overlapping VLANs, then we can define a single port on a LAN switch to be a member of all the VLANs to be linked, and then connect the router to this port via a single physical connection. Of course, we would need to configure the router to recognize all the relevant subnet ids on this port.

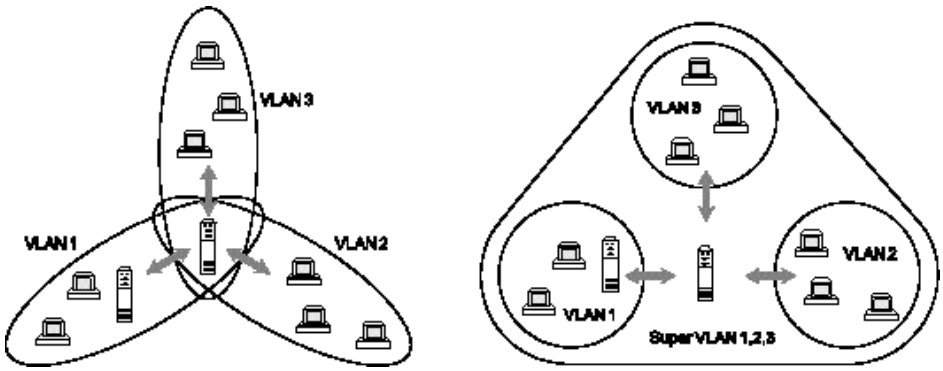

Figure 8: Flexible Overlapping and Nested Overlapping VLANs for Broadcast Control

Intelligent Broadcast Control

Some LAN switches are capable of applying a degree of intelligence to the forwarding of broadcast packets, and can therefore reduce the amount of broadcast traffic and provide security features by means of broadcast filtering. Some typical examples of intelligent broadcast filtering are:

- IP Address Caching – the switch learns IP addresses by listening to ARP broadcasts, and either responds to ARPs on behalf of a target end station, or forwards the ARP only on the port where the target end station is known to be located.
- NetBIOS Name Caching – the same technique as IP Address Caching, but applied to NetBIOS names.
- NetWare Server Discrimination – the switch learns which LAN segments have NetWare servers on them and which do not, and blocks the forwarding of SAP broadcasts to segments that have only clients attached.
- Re-circulating Explorers – the switch checks the route taken by Source Route explorer frames, and discards any such frame that has already traversed any segment attached to the switch.

The use of these techniques is a great aid to building fast, flat switched LANs by eliminating concerns about the level of broadcast traffic in the network.

Routing: Where Does It Belong?

So far, we have looked at the role of routing within the switched LAN, and we have argued that the needs of most LAN installations are best met by interconnecting all LAN segments with switches, and making minimal use of VLANs consistent with acceptable levels of broadcast traffic within the LAN. In many cases, particularly where we can deploy LAN switches with intelligent broadcast control, this will mean that we have no need to define VLANs at all. The result is a network in which all local users of IPX and non-routable protocols can freely intercommunicate via LAN switches, without ever having to traverse a router, and the only local traffic that needs to pass via a router is that between different IP subnets.

In this scenario, the amount of routing capacity needed to support the operation of the switched LAN is minimal. Often, the routers that are in place to support WAN connections can provide all the capacity that is needed. Because there is relatively little local traffic flowing to and from these routers, their location within the switched LAN is not critical.

The switched LAN architectures proposed by the router vendors, not surprisingly, look a little different from this. In general, the switched LAN is divided into many VLANs, and there is therefore a large amount of routing capacity needed to support communication between the VLANs. Where there are differences between the architectures of the router vendors, it is primarily in the location of the routing functions within the switched LAN.

Two distinctly different approaches to the location of the routing function are apparent from an examination of the various switched network architectures. One approach takes a conventional view of routing as a combination of route determination and packet forwarding, the other splits these functions and distributes them through the switched network.

The Conventional Model of Routing

The first of these two approaches involves deploying routing at one or more locations in the switched LAN, alongside the LAN switches. The routing function, which combines the elements of route determination and frame forwarding, may be carried out by conventional routers attached through one or more LAN interfaces to the switched network, or by means of router modules installed in LAN switches.

The potential problem with this approach is that traffic which has to go through a router (for example, between IP subnets) may have to be “back-hauled” through one or more LAN switches to get to the router, and that the paths to and from the router may then become bottlenecks. For example, a station connected to a LAN switch may need to communicate with a server which is on a different IP subnet, but which is connected to the same LAN switch. If there is no router connected locally to this LAN switch, then the packets that pass between these two stations will have to be sent over one or more backbone connections to reach the nearest router, and then back again.

The Distributed Routing Model

The second approach sets out to address this issue. The idea is to make every LAN switch capable of acting both as a switch and as a router, so that it can switch packets within subnets and route packets between subnets. In this approach, the switches are capable of forwarding frames at Layer 2 and at Layer 3, and the proponents of this architecture sometimes refer to this as "Multi-Layer Switching". However, if every switch were capable of doing everything that a router could do, this solution would become prohibitively expensive. So the routing function is split into two – frame forwarding and route determination – and the route determination function is implemented separately from the multi-layer switches in a centralized "route server". This simplifies the design of the multi-layer switches, and reduces the amount of memory capacity and processing power that would be needed if each switch were to do its own route determination.

In this scheme, the central route server controls all the multi-layer switches and determines, for any pair of stations that wish to communicate, what will be the path across the switched network. Furthermore, if the stations are on different subnets, the route server will determine which of the switches along the path will forward packets at Layer 2, and which will forward at Layer 3.

With the distributed routing model, the entire switched LAN becomes a kind of "virtual router". Indeed, some vendors promote their architectures in just these terms.

Distributed vs. Conventional Routing

The conventional routing scheme has the major advantages of simplicity and openness in its favor. By employing routing in a traditional way that combines Layer 3 forwarding with route determination, there is nothing new about routing for the network administrator to learn. And a clear separation between switching and routing functions allows best-of-breed solutions from different switch and router vendors to be deployed with no fear of interoperability issues.

The only real advantage that can be claimed for distributed routing over the conventional approach is that bandwidth in the LAN is better utilized if inter-subnet traffic does not have to be back-hauled across the LAN to the routers. But LAN bandwidth is relatively cheap – so back-hauling the small proportion of LAN traffic that is passing between subnets over a 100 Mbps FDDI or 155 Mbps ATM link to the routing devices does not cause any real problem of cost, performance or scalability. In any case, the amount of traffic that needs to be back-hauled can be controlled by well-planned router placement. And distributed routing has a number of serious disadvantages:

The distributed routing scheme is far more complex. It involves the introduction of new kinds of route determination algorithms to handle the multi-layer switch concept, and proprietary protocols for distributing routing information to the multi-layer switches. It will take years for these new algorithms and protocols to mature, and there is considerable technical risk in the meantime.

Trouble-shooting is likely to move into a new dimension of complexity. For example, for any given session between two end stations on different subnets that passes through multiple switches, it may be difficult to determine which switches are forwarding at Layer 2 and which at Layer 3. This has obvious implications for ease of diagnosing problems.

Distributed routing involves proprietary protocols and is therefore a single vendor solution. Because the distributed routing model requires the introduction of new and proprietary protocols between route servers and multi-layer switches, this approach is effectively a single-vendor solution. While proprietary solutions may occasionally be desirable where they deliver far higher functionality than standards-based approaches, in this case the downside of single vendor lock-in does not look like a good trade-off for the supposed advantages of distributed routing.

Deploying Conventional Routing

Given these disadvantages of distributed routing, the conventional approach to routing is to be preferred. But this does not imply necessarily that we need to employ conventional routers to perform the routing function. What it does imply, though, is that the routing function shall combine route determination with Layer 3 forwarding in a conventional manner.

If we choose to use conventional routers, we can connect them to the switched LAN by means of one or more standard LAN interfaces. For example, if we determine that we need a total routing throughput of 20 Mbps to handle the local inter-subnet traffic, then we could connect a router to one of the LAN switches by means of four 10 Mbps Ethernet interfaces. (To get 20 Mbps total throughput, we must provide 20 Mbps into and 20 Mbps out of the router).

For higher throughput at lower cost, we could adopt the “one-legged” router approach. Here we use a conventional router with a single very high speed interface to the switched LAN, typically 155 Mbps ATM. Since ATM is a full duplex technology, such an interface can support a full routing throughput of 150 Mbps or more – provided, of course, that the router can forward at this rate.

We can take this concept one step further, by integrating the routing function directly into a LAN switch or ATM switch. Here we can save the cost of the physical interfaces needed to connect the one-legged router by attaching the routing function directly to the LAN switching or ATM cell switching fabric, to give us the greatest possible cost efficiency.

Although this approach sounds as though it is the same as the “Multi-Layer Switch”, there is one key difference: the Multi-Layer Switch receives its Layer 3 forwarding tables from a central route server, while the integrated routing function includes the full route determination capability. Instead of integrating part of the routing function into every LAN switch, what we are doing is integrating all of the routing function into some of the LAN switches.

ATM in the Switched LAN

In the previous section on the deployment of routing in the switched LAN, we touched for the first time on the subject of a new networking technology that is rapidly becoming of great relevance to the switched LAN: Asynchronous Transfer Mode (ATM).

There is not space here to cover the subject of ATM in any detail: the reader seeking more information is referred to white papers published by Madge Networks on ATM at the Desktop and LAN Emulation over ATM, as well as numerous widely available books and articles.

For some network planners, ATM has immediate relevance, whereas for others its importance may lie beyond the current planning horizon. Nevertheless, all network planners responsible for the creation of large switched LANs should take into account the effects of possible future deployment of ATM technology in the network, and its implications for the mix of switching and routing within the LAN.

ATM is a switched networking technology which uses small, fixed length cells to transport information rather than the variable length packets we are used to with LAN technologies such as Ethernet. Also, unlike Ethernet and Token Ring, ATM is connection-oriented. This means that a connection has to be set up between two end-points in the network, rather like dialing a phone call, before any data can be sent. Like LAN switching, ATM's advantages include very high speed, low latency, and excellent price/performance, but ATM offers further benefits with its ability to carry multiple different kinds of traffic including voice and video alongside data, with guaranteed quality of service appropriate to each type of traffic.

ATM is starting to be deployed quite widely in LANs, both as a high capacity LAN backbone, and to support demanding desktop applications.

LAN Emulation over ATM

Powerful though it is as a networking technology, ATM would have no interest for the network planner if it were not able easily to be integrated with existing LAN technologies, and support existing LAN applications. Since ATM is based on a connection-oriented cell transport, it is not obvious how it can be made to fit in a world that is dominated by the connectionless frame transport provided by Ethernet, Token Ring and FDDI.

The most popular solution to this problem currently is known as LAN Emulation over ATM. In this scheme, Ethernet and Token Ring frames are transported over ATM by segmenting complete frames, including source and destination MAC address headers, into streams of ATM cells, and mapping destination MAC addresses into the Virtual Circuit numbers that are used to address the ATM cells. Since this technique makes an ATM network into a transport for Ethernet and Token Ring frames, it can be treated as if it were a real Ethernet or Token Ring network (albeit a very fast one) by LAN switches and routers attached to it.

Furthermore, a single ATM network may support multiple Ethernet and Token Ring broadcast domains, each of which can be regarded as a separate emulated LAN, and these emulated LANs can be thought of as equivalent to Virtual LANs in the frame-switching world.

With the ATM network configured to support one or more emulated LANs, we can connect existing physical LANs to it by means of either switches or routers. LAN switches that support an ATM interface with LAN Emulation (LANE) will simply treat this port as a very fast Ethernet or Token Ring connection, and will switch frames to it according to the addressing or source routing information in the frames. This kind of switch will be able to make connections between multiple VLANs on the Ethernet or Token Ring side with multiple Emulated LANs (ELANs) on the ATM side. Generally, each VLAN will connect to one ELAN, so that the broadcast domain created by the VLAN across the frame-switched network now extends via the ELAN into the ATM world.

Equally, routers with ATM connections that support LANE could be used to connect existing LANs to an ATM network, although here, the speed advantages of the switched network are likely to be compromised by the performance limitations of the router.

When considering how routing should be deployed in a switched LAN that incorporates ATM with LAN Emulation, all we need to do is to think of the Emulated LAN as an extension of the physical switched Ethernet or Token Ring LAN. Following the recommendation made in the previous section, the routing function is best deployed in a centralized fashion at a handful of points in the switched LAN, and for best performance it should be attached directly to the high speed switched ATM backbone – in the form of a “one-legged router” with ATM interface, or integrated into the ATM switches themselves.

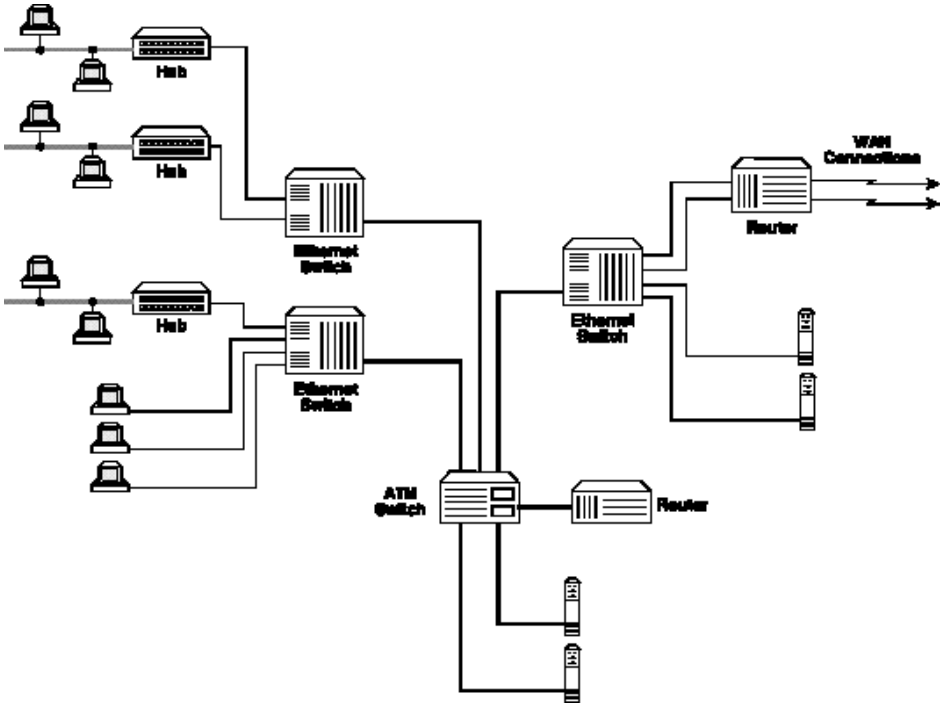

Figure 9: ATM Backbone with One-Legged Router

Multi Protocol over ATM

LAN Emulation provides a simple and effective solution to the problem of integrating ATM with existing switched LANs, allowing switched connections be built between Ethernet and Token Ring networks and ATM, with support for existing LAN applications regardless of the network layer protocol they use. However, LANE still suffers from the same limitation we see in switched LANs: that we cannot communicate between two stations with different IP subnet addresses without passing through at least one router.

Work is currently in progress on a new standard for ATM networking which aims to address this limitation. The work is initially focused on IP, but aims eventually to provide solutions for other widely-used protocols – hence the name Multi Protocol over ATM (MPOA). With MPOA, IP frames are carried directly in ATM cells without the overhead of Ethernet or Token Ring MAC headers imposed by LANE.

A declared aim of MPOA is that any two stations connected by a switched network should be able to communicate without the need for a router to forward packets between them, regardless of their IP address. The fact that the work on MPOA to achieve this aim is proceeding with the enthusiastic support of major router vendors is a clear vindication of the position taken by this paper: that it is far preferable for end stations in a network to intercommunicate directly via switched connections rather than via a router.

At the time of writing, there are still many detailed aspects of MPOA to be worked out, but the broad principles are fairly well established. The implementation of MPOA involves changes to the behavior of end station IP protocol stacks, as well as the introduction of a new kind of distributed route server function.

Current end station IP protocol stacks compare their own subnet address with that of their intended target destination. If they are the same, the station concludes that its target is on the same subnet as itself, and it then uses the broadcast Address Resolution Protocol to find the MAC address of the target. If not, the station concludes that the target is on a different subnet, which is only reachable via a router, and it therefore initiates communication with the target by sending packets to the router.

With MPOA, the end station does not automatically send packets to a router in order to communicate with a station that is not on its local subnet. Instead, it queries a route server which, in cooperation with other route servers in the network, holds reachability information for all stations on the switched network. If it is possible to reach the target station via a switched connection, the route server will respond with the appropriate address information for the target and the requesting station can then initiate communications with the target directly over the switched network.

While the direction of MPOA is clear, there is much work that remains to be completed, both to deal with the operation of mixed Ethernet, Token Ring and ATM networks, and to develop the protocols used by the route servers to maintain distributed reachability information. In the future, MPOA holds the promise for even more effective exploitation of switching for fast, efficient networking, but for the time being, the established standard for LAN Emulation provides the only realistic solution for integrating ATM into existing switched LANs.

Conclusion

The introduction of switching changes the fundamental rules of LAN design. Deploying this new technology wisely means taking full advantage of the virtues of speed and simplicity that are inherent to switching. Using switching merely to virtualize LANs around routers misses the point: switching is the new paradigm of networking, and routing, though essential to the proper functioning of a practical network, is no longer the axis around which all else revolves.

To summarize the main recommendations of this paper about the design of switched LANs:

- Connect all LAN segments with LAN switches, so that any station can reach any other station without there being a router in the physical path.
- Connect routers only to LAN switches, and to WAN links.
- Consider the traffic flows of all protocol types, to make sure that IP addressing limitations aren't forcing all other protocols to pass unnecessarily through routers.
- Use VLANs sparingly, both to minimize the burden of administering VLAN configuration, and to minimize the amount of costly router capacity needed to pass packets between VLANs.
- Choose LAN switches that support overlapping VLANs or intelligent broadcast control features to enable larger, flatter switching domains to be constructed without broadcast problems.

Madge Networks

Americas
2310 North First Street
San Jose
CA 95131/1011
United States
Tel +1 408 955 0700
Fax +1 408 955 0970
http://www.madge.com

Asia, Australia & New Zealand
12 / F Li Po Chun Chambers
189 Des Voeux Road
Central
Hong Kong
Tel +852 2593 9888
Fax +852 2519 8022
http://www.madge.com

Europe, Middle East & Africa
Knaves Beech Business Park
Loudwater, High Wycombe
Bucks HP10 9QZ
England
Tel +44 1628 858000
Fax +44 1628 858011
http://www.madge.com

Japan
Mita NN Building
1-23, Shiba 4-chome
Minato-Ku, Tokyo 108
Japan
Tel +81 3 5232 3281
Fax +81 3 5232 3208
http://www.madge.com

980-727-01 3/96