About this Guide

This document provides information about designing networks that transport SNA protocols over Frame Relay using an IP backbone.

Often, discussions about SNA over Frame Relay lead to SNA-only solutions such as IBM's Boundary Network Node (BNN) or Boundary Access Node (BAN). The relative simplicity of BNN and BAN solutions are attractive to some customers. These customers have migrated toward BNN and BAN to transport SNA. However, the vast majority of Cisco customers use DLSw+ to not only transport SNA, but integrate SNA into an IP backbone that uses Frame Relay.

Frame Relay is a Layer 2 technology. IP is a Layer 3 technology. TCP encompasses Layers 4 through 6 technologies. SNA is the application data. In the case of SNA over Frame Relay, the application data (SNA) is a suite of protocols that have their own architecture encompassing Layers 2 through 7. Cisco has proven that these two diverse protocol suites can peacefully converge over a common enterprise network using leased lines to access remote sites. Frame Relay services to remote locations are decreasing in price and increasingly ubiquitous. These factors are driving decision-makers to replace existing wide-area leased-line infrastructures with equivalent Frame Relay services. However, the lower cost of Frame Relay comes at a price of increased complexity that must be addressed to maintain existing user service levels. To specifically address DLSw+ over Frame Relay, this guide was written.

This guide does not include topics presented in other related design guides, although there is some minor overlap. Non-SNA topics on Frame Relay are presented in *Frame Relay Design Guide*. DLSw+ topics are presented in *DLSw+ Design and Implementation Guide*. APPN topics are presented in *APPN Design and Implementation Guide*.

Cisco documentation and additional literature are available in a CD-ROM package, which ships with your product. The Documentation CD-ROM, a member of the Cisco Connection Family, is updated monthly. Therefore, it might be more current than printed documentation. To order additional copies of the Documentation CD-ROM, contact your local sales representative or call customer service. The CD-ROM package is available as a single package or as an annual subscription. You can also access Cisco documentation on the World Wide Web at http://www.cisco.com, http://www-china.cisco.com, or http://www-europe.cisco.com.

If you are reading Cisco product documentation on the World Wide Web, you can submit comments electronically. Click **Feedback** in the toolbar, select **Documentation**, and click **Enter the feedback form**. After you complete the form, click **Submit** to send it to Cisco. We appreciate your comments.

Intended Audience

This document is for anyone who wants to learn more about Cisco's SNA internetworking over Frame Relay solutions. A technology overview, including a brief historical perspective, is included to provide background information and a sound technical understanding of SNA and Frame Relay. Special emphasis is placed on issues related to maintaining SNA response times, especially the causes that degrade SNA response times and the techniques to improve them. Also included is a section on the performance characteristics of the various encapsulation methodologies for Frame Relay (performance testing for BNN and BAN has not been done as of this writing). Although DLSw+ is a major part of this design guide, some features (for example, backup and redundancy) are not discussed, because they are covered in detail in *DLSw+ Design and Implementation Guide*.

Examples of key configuration commands are shown to aid in understanding a particular configuration. However, this document does not contain the exact and complete configurations. This information is available and regularly updated in Cisco Connection Online (CCO) and in the Cisco product documentation. CCO is Cisco's primary, real-time support system and is accessible at the World Wide Web address http://www.cisco.com.

Document Structure

This document contains the following chapters and appendixes:

- Technology Overview—Provides an overview of the technologies involved in SNA internetworking over Frame Relay.
- Maintaining SNA Response Times—Discusses the most prevalent problems that degrade response times in an SNA internetworking over Frame Relay scenario and techniques to resolve them.
- Router Performance Considerations—Discusses the relative performance characteristics of each DLSw+ encapsulation method available for Frame Relay.
- Glossary—Provides a list of terms and acronyms that are relevant to this guide.
- References—Provides a list of related documents.

Cisco Connection Online

Cisco Connection Online (CCO) is Cisco Systems' primary, real-time support channel. Maintenance customers and partners can self-register on CCO to obtain additional information and services.

Available 24 hours a day, 7 days a week, CCO provides a wealth of standard and value-added services to Cisco's customers and business partners. CCO services include product information, product documentation, software updates, release notes, technical tips, the Bug Navigator, configuration notes, brochures, descriptions of service offerings, and download access to public and authorized files.

CCO serves a wide variety of users through two interfaces that are updated and enhanced simultaneously: a character-based version and a multimedia version that resides on the World Wide Web. The character-based CCO supports Zmodem, Kermit, Xmodem, FTP, and Internet e-mail, and it is excellent for quick access to information over lower bandwidths. The WWW version of CCO provides richly formatted documents with photographs, figures, graphics, and video, as well as hyperlinks to related information.

You can access CCO in the following ways:

- WWW: http://www.cisco.com
- WWW: http://www-europe.cisco.com
- WWW: http://www-china.cisco.com
- Telnet: cco.cisco.com

- Modem: From North America, 408 526-8070; from Europe, 33 1 64 46 40 82. Use the following terminal settings: VT100 emulation; databits: 8; parity: none; stop bits: 1; and connection rates up to 28.8 kbps.

For a copy of CCO's Frequently Asked Questions (FAQ), contact cco-help@cisco.com. For additional information, contact cco-team@cisco.com.

Note: If you are a network administrator and need personal technical assistance with a Cisco product that is under warranty or covered by a maintenance contract, contact Cisco's Technical Assistance Center (TAC) at 800 553-2447, 408 526-7209, or tac@cisco.com. To obtain general information about Cisco Systems, Cisco products, or upgrades, contact 800 553-6387, 408 526-7208, or cs-rep@cisco.com.

Technology Overview

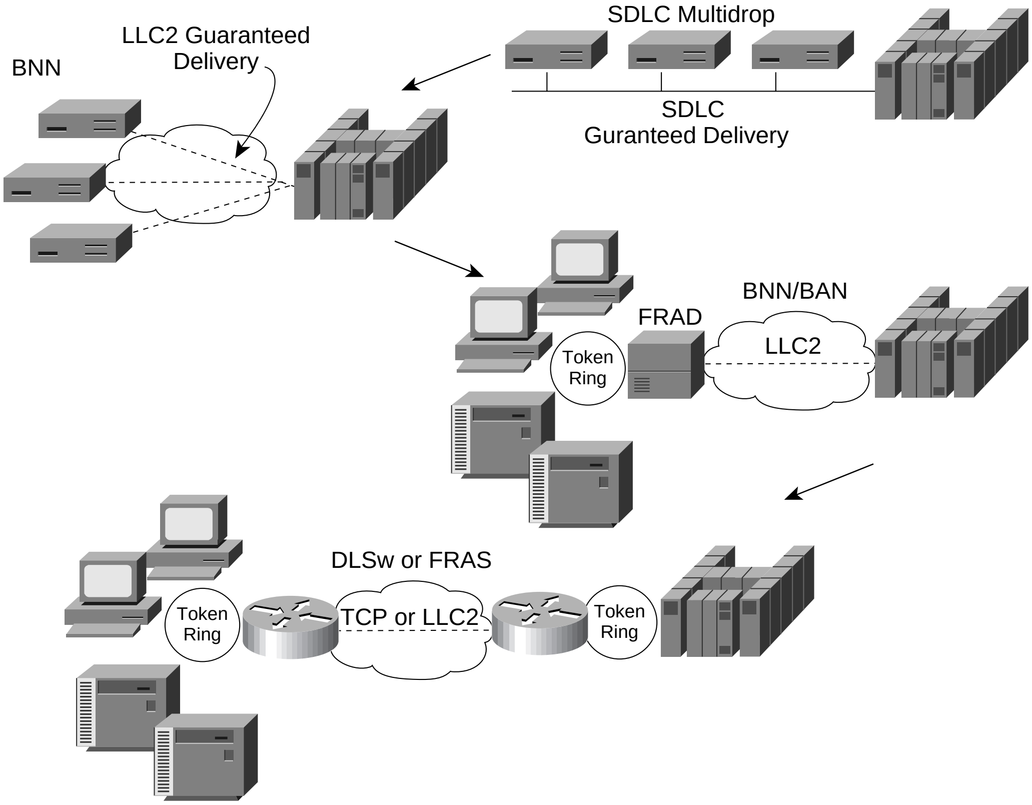
This chapter provides a historical perspective that shows how the SNA technology evolved to the state it is at today. Frame Relay standards and how they apply to SNA are also presented. Last, frame formats and the flow of data are discussed.

Historical Perspective

The reason that technology to carry SNA traffic over Frame Relay exists is to maintain corporate investments in IBM mainframe applications. For many corporations, a primary requirement is to never change mainframe applications. This requirement is so important that all other aspects of the networking enterprise must accommodate it. This is pervasive throughout the industry.

Figure 1-1 provides an overview of the SNA internetworking technology evolution. In the beginning there was the classic multidrop Synchronous Data Link Control (SDLC) protocol scenario. SDLC was the dominant link-layer protocol. SNA relied on SDLC to provide a guaranteed delivery service, thus avoiding the need to address mundane tasks such as monitoring line quality. In Europe and Asia, in addition to SDLC, X.25 was a very predominant transport mechanism for SNA. Like SDLC, X.25 provided the guaranteed delivery service required by SNA. Qualified Logical Link Control (QLLC) was implemented on Network Control Program (NCP) to make Virtual Telecommunications Access Method (VTAM) think that SDLC was running on the front-end processor (FEP). QLLC essentially performed SDLC-like functions over X.25. IBM provided this function in a product called X.25 NCP Packet Switching Interface (NPSI).

Figure 1-1 Historical Evolution of SNA Internetworking Technology



As technology progressed, more-reliable digital technology paved the way for new, higher-speed packet services that assumed the presence of reliable links. From these fast-packet services, frame-switching services emerged and eventually became Frame Relay as we see it today. The presence of reliable links means 99.999 percent network reliability, which is fine for datagram-based protocols (for example, IP, IPX, and AppleTalk) that have higher-layer subsystems handling guaranteed delivery. This level of reliability is not acceptable to SNA. In addition, the frame-based services needed to incorporate statistical multiplexing functionality, which was popular at the time. This feature led to what is referred to as over-subscription. By over-subscribing, networks exposed themselves to the possibility that periods of congestion could occur. Congestion results in frame loss, which is unacceptable to SNA. However, for IBM customers, the low-cost, frame-based networking service was acceptable.

To allow IBM customers to take advantage of these new low-cost, frame-switching services, a guaranteed delivery link-layer service was needed to replace SDLC and QLLC. To replace these protocols, IBM found an existing and proven LAN-based protocol to do the job. Logical Link Control, type 2 (LLC2) was chosen as the protocol to run over Frame Relay. Thus, LLC2 appeared on IBM remote cluster controllers and on NCP when it was used over Frame Relay. On NCP the feature was called BNN. BNN was designed to allow a single IBM 3270 cluster controller residing on a Frame Relay connection (virtual circuit) access to the mainframe. BNN masked any remote access peculiarities related to Frame Relay from VTAM.

BNN marked a significant event in the industry because mainframe access was now conducted over Frame Relay. Before long, many of the existing X.25 packet assembler/disassembler (PAD) vendors saw market potential for a similar device for SNA Frame Relay environments. These devices were called Frame Relay access devices

(FRADs). The idea was to take a number of legacy SNA devices running, for example, legacy SDLC and connect them to a FRAD to perform a protocol conversion to the BNN formats on Frame Relay. As a result, the SNA devices appeared to NCP to be IBM 3270 cluster controllers at the remote end of the Frame Relay connections. The next step was to connect an entire LAN to the FRAD and perform the same conversion process. BNN assumes only a single device for each Frame Relay connection. To accommodate multiple devices, one had to use multiple connections (virtual circuits) or some sophisticated physical unit (PU) concentration that presented the mainframe with a single SNA device from multiple collapsed SNA devices. These techniques were more costly and complex, but they did work.

Multiple devices can be supported on a single virtual circuit using a technique called service access point (SAP) multiplexing. This multiplexing is accomplished by manually mapping the legacy Layer 2 device address (media access control [MAC] or SDLC station address) to the LLC2 SAP address on the Frame Relay WAN. The FRAD devices and the FEP can uniquely identify the legacy devices from a destination service access point/source service access point (DSAP/SSAP) pair appearing on the virtual circuit. By varying the LLC2 SAP address, multiple devices can be represented. The administrator configures a map from the LAN, SDLC, or X.25/QLLC device address to a unique LLC2 DSAP/SSAP pair. This configuration worked for small environments, but something more encompassing was needed for large enterprises with many LAN devices. (For a description of the BNN frame format, see the "SNA Internetworking Frame Formats" section.)

BAN was designed from the beginning with the FRAD capabilities in mind. Using the RFC 1490 bridged format, BAN added fields to the frame formats that allow MAC addresses to appear on the WAN along with the LLC2 SAP addresses. This additional capability was the most significant functional change introduced by BAN. The number of devices supported was virtually limitless and, because MAC addresses were passed through, no additional configuration was required on the FRAD to support multiple LAN devices. BAN did required changes to be made to NCP because new configuration would be required. However, now the industry had two major methods of accessing the mainframe over Frame Relay. (For a description of the BAN frame format, see the "SNA Internetworking Frame Formats" section.)

Today, there are more than two methods for accessing the mainframe over Frame Relay. While development was taking place on cluster controllers, NCP, and FRAD devices, similar efforts were afoot to allow SNA protocols to operate over IP-based networks. Like Frame Relay, IP is a best-effort delivery service. TCP/IP was used to furnish the guaranteed delivery requirement for SNA in these environments. Initially, Cisco delivered remote source-route bridging (RSRB), which was tremendously successful in the market. IBM, largely rejecting the SNA internetworking market at that time, had limited success with DLSw based on an earlier RFC 1434 standard implementation. The standard had so many problems that it was eventually abandoned. In these environments using DLSw, NCP never came into the picture. These bridging and switching techniques were to be implemented on a different type of general-purpose networking devices—routers.

Cisco, IBM, and the rest of the internetworking industry eventually embraced the DLSw standard and ratified RFC 1795. In a parallel effort, FRAD vendors implemented the host portion of BNN and BAN on their devices, eliminating the need to upgrade the FEPs to run BNN or BAN. Routers that route IP and carry SNA over IP on Frame Relay work well. FRADs carrying SNA over BNN or BAN and routing IP work well. As a result, there was a convergence of routing products and FRAD products and their respective markets.

Some FRAD vendors and companies implementing SNA-only technologies claim that there is a technological difference between the methods they use to carry SNA over Frame Relay and the methods that routers use. Technically, there is no difference. A Frame Relay switch, which connects to both routers and FRADs, certainly cannot distinguish between the two. The only difference lies in the background of the various vendors and the markets in which those vendors are comfortable. FRADs have evolved from vendors with backgrounds in circuit-based transmission systems or Layer 2 technologies. These vendors are more familiar with technologies like BNN, BAN, voice, and X.25. Router products have evolved from vendors with backgrounds in networking-based technologies that are media independent and function at Layer 3 and higher. FRAD marketing programs, sometimes backed by paid industry consultants, that attempt to position their products as technically superior. These issues are clearly non-technical.

Cisco, in order to remain technology-neutral, supports all of the FRAD and DLSw technologies for SNA. Because of the combination of Cisco's IP routing history, its huge customer base with an existing IP infrastructure, and the attraction of an all-IP backbone, DLSw+ has emerged as the dominant technology in Frame Relay-based networks carrying SNA protocols.

Frame Relay Standards

There are numerous Frame Relay standards. In addition to specifying frame formats and protocols, one of the primary objectives of Frame Relay standardization is to agree on a means of identifying multiple protocols over a single Frame Relay virtual circuit to allow interworking. The goal is to administer a common set of protocol IDs. The list of organizations employed to develop the common protocol IDs is impressive: ITU-T, ANSI, AIW, FRF, IETF, ISO, and the former CCITT. Adding to the confusion are multiple standards that basically accomplish the same thing with few end-user perceived differences. The following Frame Relay standards are considered most relevant:

- ITU-T Q.922—Integrated Services Digital Network (ISDN) data link layer specification for frame-mode bearer services
- ITU-T Q.933—ISDN digital subscriber DSS 1 signaling specifications for the frame mode switched and permanent virtual connection control and status monitoring
- ANSI T1.617a Annex F—DSS1 signaling specification for Frame Relay bearer service
- ISO 9577—Information technology protocol identification in the network layer

Most of the standards above enumerate a set of protocol IDs referred to as network layer protocol IDs. From these works, the following two major Frame Relay internetworking specifications have emerged for SNA:

- RFC 1490—Multiprotocol interconnect over Frame Relay
- FRF 3.1—Frame Relay multiprotocol implementation agreement

RFC 1490 covers encapsulation methods for bridging and routing multiprotocol internetworking traffic over Frame Relay. FRF 3.1 covers user-specified network layer protocol IDs for SNA not found in RFC 1490, for example NetBIOS and High Performance Routing (HPR).

For more information on standards documents, see the References appendix.

SNA Internetworking Frame Formats

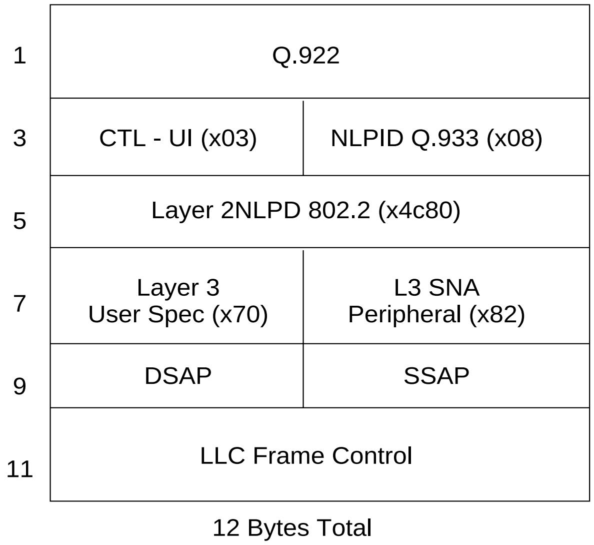
Most networking equipment today uses RFC 1490 or FRF 3.1 encapsulation methods when carrying SNA protocols over Frame Relay. A common concern about SNA internetworking is encapsulation overhead. There is little technical merit for this concern and its roots can be traced to marketing claims. The following frame formats are discussed in this section:

- RFC 1490 Routed Frame Format (BNN)
- RFC 1490 Bridged Frame Format (BAN)
- DLSw+ Lite

RFC 1490 Routed Format

BNN uses the routed frame format for transporting SNA over Frame Relay although legacy SNA PU 2.0 protocols are not routable. Advanced Peer-to-Peer Networking (APPN), HPR, and SNA format indicator 4 (FID 4) subarea protocols are routable so there is some justification for the terminology. In addition, the MAC addresses are removed, which meets at least one criteria of a routed protocol. RFC 1490 defines the format of a routed frame. FRF 3.1 is used to specify the user-defined Layer 2 and Layer 3 protocol IDs commonly used today. Figure 1-2 is a diagram of the BNN frame format.

Figure 1-2 RFC 1490 Routed Frame Format (BNN)

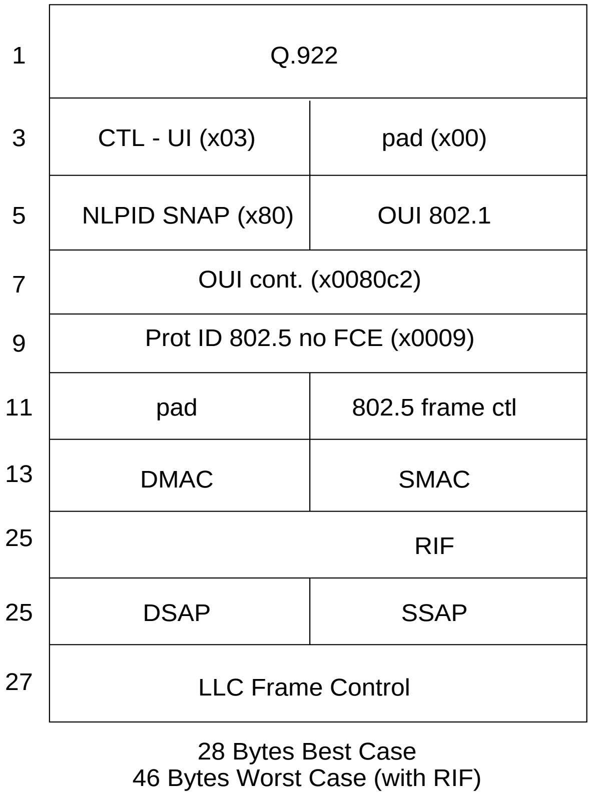


BNN, with only 12 bytes of overhead, has the lowest overhead compared to BAN, DLSw+, and DLSw+ Lite. BNN uses a network layer protocol ID set to a value indicating that CCITT Q.933 formats are used (x08). This allows an additional 4 bytes of protocol identification for Layer 2 and Layer 3. The Layer 2 protocol is set to 802.2 or LLC2 (x4c80), and the Layer 3 protocol is set to user specified (x70) and SNA peripheral (x82). These codes-points for the Layer 3 network layer protocol ID are documented in FRF 3.1.

RFC 1490 Bridged Format (BAN)

BAN uses the bridged frame format for transporting SNA over Frame Relay, although IBM states that BAN is not bridging. This statement is a marketing ploy and has little technical merit. BAN, is at best a highly filtered bridging technology. Figure 1-3 is a diagram of the BAN frame format.

Figure 1-3 BAN Frame Format



1	Q.922	
3	CTL - UI (x03)	pad (x00)
5	NLPID SNAP (x80)	OUI 802.1
7	OUI cont. (x0080c2)	
9	Prot ID 802.5 no FCE (x0009)	
11	pad	802.5 frame ctl
13	DMAC	SMAC
25	RIF	
25	DSAP	SSAP
27	LLC Frame Control	

28 Bytes Best Case
46 Bytes Worst Case (with RIF)

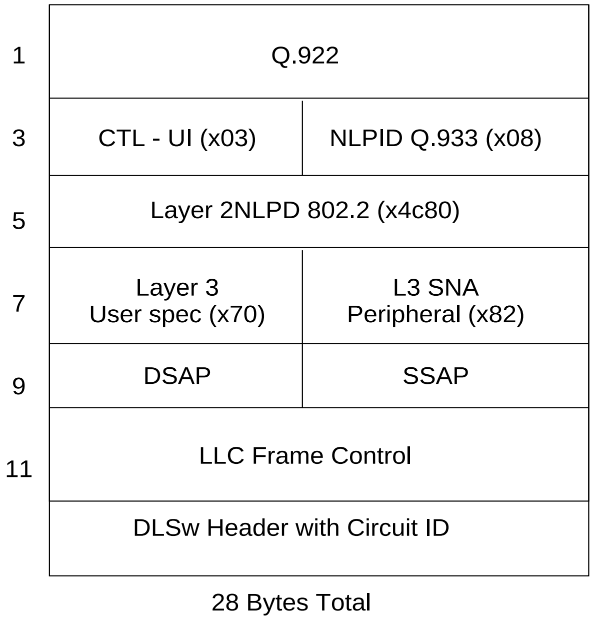
The encapsulation overhead for BAN is usually 28 bytes unless there is a Routing Information Field (RIF) present. Cisco's implementation terminates the RIF and strips it from the frame before placing the frame on the WAN. Using the protocol IDs specified in RFC 1490, the network layer protocol ID is set to Subnetwork Access Protocol (SNAP) (x80). Within the SNAP header the Organizational Unique Identifier (OUI) is set to a value of 0x0080c2, indicating bridged format, and a PID of x0009, indicating 802.5. MAC addresses are also included in the frame format. The addition of MAC addresses gives BAN a virtually unlimited capacity to support bridged devices.

DLSw+ Lite

DLSw+ Lite is a special consideration because it does not fall into any ratified standard. Cisco offered DLSw+ Lite to the standards organizations, but, because of a lack of customer demand beyond Cisco, there was little interest from other vendors. DLSw+ Lite uses the RFC 1490 routed format indicating user-specified IDs for Layer 2 and Layer 3 using FRF 3.1. DLSw+ Lite uses the following routed frame format: network layer protocol ID set to indicate CCITT Q.933 protocol ID (x08), 802.2 or LLC2 Layer 2 protocol (x4c80), user-specified Layer 3 protocol (x70), and SNA DLSw (x81).

The encapsulation overhead of DLSw+ Lite is 28 bytes, which is the same as BAN. Figure 1-4 shows the DLSw+ Lite frame format. DLSw+ Lite is the Frame Relay equivalent of DLSw+ direct.

Figure 1-4 DLSw+ Lite Frame Format



Data Flow

The major protocols that the network designer needs to be concerned with are LLC2, NetBIOS, and HPR. NetBIOS uses LLC2, but it also uses LLC1 frames or datagrams to broadcast information about resources and to locate NetBIOS resources. Both SNA (including APPN) and NetBIOS use explorer broadcasts or multicast frames (called functional addresses in Token Ring) to locate resources. After resources are located, both protocols use LLC2 for session establishment and information transfer. HPR uses only LLC1 datagrams (unnumbered information [UI] frames). Like TCP/IP, HPR relies on higher-layer protocols, such as Rapid Transport Protocol (RTP), to guarantee delivery.

On the LAN there are basically three major categories of frame types: explorer frames, information frames, and datagrams. Explorer frames can be broken down into local explorers, all-routes explorers, single-route explorers, and directed explorers. Information frames can be broken down into supervisory frames, which are used to initiate and synchronize link-layer establishment, and information frames (I-frames), which carry data. Datagrams, or UI frames, are sent without any guarantee of delivery.

Broadcast and multicast explorers are used to locate resources. They are forwarded by bridging devices. As explorers are forwarded, the bridging devices add the input ring that the explorer came from and its bridge ID to the RIF. The bridging device then forwards the explorer through each Token Ring port configured as part of the bridging domain. The bridging domain information is derived from the **source-bridge active** command on the Token Ring interfaces or the **bridge-group** command on Ethernet interfaces. Each bridge checks the existing RIF to make sure that an explorer is not put on the same ring twice to avoid a bridging loop. When the explorer reaches the target ring, the resource or target station is contacted. The target station saves the RIF and responds directly to the originator by reversing the RIF. The originator saves the RIF on the response as well. Subsequent communication takes place using the learned RIF. Traditionally, source-route bridges do not get involved with storing ring routing information. The RIF is stored at the source end stations. Thus the process is called source-route bridging (SRB). The router facilitates value-add features such as proxy explorer, explorer firewalls, and multiring routing. A RIF cache is used on the router to implement these features, but these features are not required to perform basic source-route bridge functions.

As mentioned before, there are four types of explorers: local explorer, single-route broadcast explorer, all-routes broadcast explorer, and directed explorer. An originating end station first sends a local explorer to locate resources locally on a ring. A local explorer is never picked up by a bridging device. When the originating device is convinced that the resource cannot be found locally, it sets the source-route indicator bit in the SMAC header. Setting this bit signals locally attached bridges to pick up the frame and forward it. However, in fully connected redundant ring topologies with multiple paths to resources, this forwarding of the all-routes broadcast explorers can create broadcast storms. To reduce the impact of this phenomenon, single-route broadcast explorer frames were introduced. SNA sends all-routes broadcast and NETBIOS sends single-route broadcasts. Bridging devices forward single-route broadcast frames on a single path as determined by a spanning-tree protocol run on the bridge. Another form of explorer frame is the directed explorer. The directed explorer is commonly used by DLSw+. It contains a valid RIF and is sent directly to the resource for verification of its presence.

DLSw+ Explorer Handling

DLSw+ picks up all-routes and single-route explorers and unconditionally converts them to single-route explorers. This default behavior can be changed using a configuration option. DLSw+ also terminates the RIF by stripping it from the frame and storing it in the reachability cache, which is indexed by target resource identifier (MAC address or NetBIOS name) for future use. Stripping the RIF and using the reachability cache allows DLSw+ to expand source-route bridged networks beyond the maximum architectural limit of seven. Going beyond seven hops is not recommended, but the ability is present.

DLSw+ uses a protocol called Switch-to-Switch Protocol (SSP) between peers, which is specified in RFC 1795. SSP has primitives that are equivalent to Token Ring LLC or NetBIOS protocol data units. As a result, a protocol conversion takes place from LLC2 to SSP. For example, explorers are converted to SSP CANUREACH frames. SSP is encapsulated over the WAN (using, for example, TCP/IP, Fast-Sequenced Transport [FST], direct High-Level Data Link Control [HDLC], or LLC2).

Session Establishment and Data Transfer

Once target resources are contacted, the next step is to establish an LLC2 session with the target and then maintain that session, guarantee the delivery and sequence of frames, and maintain flow control during periods of congestion. Meeting these requirements, while sending data over a low-speed WAN has resulted in numerous features with various tradeoffs.

For example, to maintain the session, a technique called local acknowledgment (LACK) was developed. LACK is similar to a spoofing technique used in satellite networks for many years. LACK terminates the LLC2 session and locally maintains sequencing, retransmission, and flow control. However, LACK has drawbacks. It adds a measurable amount of latency that can affect high-traffic networks. In addition, not all WANs are low speed, so LACK is not necessary in all circumstances. To address these drawbacks, LACK can be disabled and LLC2 protocols can be supported across the WAN in a passthrough mode where the end stations once again resume the responsibilities of sequencing, retransmission, and flow control.

With either DLSw+ or FRAS running, LACK works in the following manner: explorers are passed through and supervisory frames are locally acknowledged, but synchronization of the session establishment does not occur until the remote station responds. I-frames are locally acknowledged. Flow control is handled locally.

It is important to note that even though data has been acknowledged, it has not been delivered. Acknowledgment means that data has been queued for delivery to a guaranteed delivery subsystem such as TCP/IP or LLC2 on the WAN. If the WAN session is lost, then all locally acknowledged LAN sessions that rely on that WAN session must disconnect because the data may have not been delivered to the remote location. There is no way to know for sure if data has been delivered so the only alternative is to disconnect the session.

In passthrough mode, all LLC frames pass through intermediate stations to the remote target stations. All the sequencing, retransmission, and flow control is handled by the end stations. If the intermediate WAN session is lost, then the end station LLC2 stack detects the failure (by timing out retransmissions) and implements recovery procedures. Passthrough modes are rare in Frame Relay environments because of the low speeds that are often present.

DLSw+ Session Establishment and Data Transfer

DLSw+ continuously converts from LLC2 on the LAN to SSP formats on the WAN. This conversion takes place regardless of the WAN encapsulation method. Thus, SSP is encapsulated in TCP/IP, IP (that is FST), direct HDLC, or LLC2.

DLSw+ has multiple encapsulation methods with varying levels of encapsulation overhead. However, in most cases the overhead is negligible and can be safely omitted as a design consideration. A discussion of each encapsulation method is beyond the scope of this guide. Please see *DLSw+ Design and Implementation Guide* for more information about encapsulation.

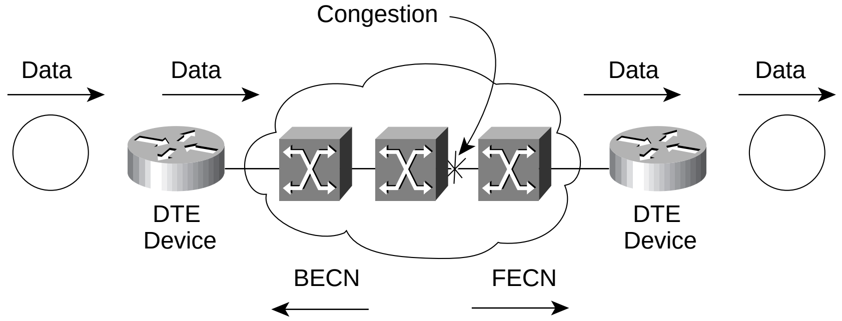
Congestion Management

Frame Relay networks attracted attention because of their ability to more efficiently use shared backbone resources such as bandwidth. Frame Relay surpasses the traditional time-division multiplexing (TDM) functionality by allocating unused bandwidth to a connection that needs more bandwidth. This capability is the source of Frame Relay's success, but it also the source of many of its problems. To accommodate this ability, Frame Relay network designers have the additional task of engineering traffic in a process called provisioning. Now, instead of connections having dedicated bandwidth, a connection is assigned varying amounts of bandwidth, in a well-defined range, depending on the traffic conditions. Simply put, if traffic is light, connections get more bandwidth; if traffic is heavy, then connections get less bandwidth.

Explicit Congestion Notification

Two bits in the Frame Relay header are reserved to notify the directly connected end systems about the current traffic conditions. The backward explicit congestion notification (BECN) bit and the forward explicit congestion notification (FECN) bit send congestion indication information (see Figure 1-5). Though simplistic in principle, FECN and BECN mechanisms are a source of infinite confusion. The confusion may be because the terminology chosen is from the perspective of the network and not the perspective of the user. When congestion is detected anywhere along the path in a Frame Relay virtual circuit, the traffic heading in the same direction (forward traffic) has the FECN bit set on it. Traffic heading in the opposite direction (backward) has the BECN bit set on it. How the individual Frame Relay networks determine that congestion is present is specific to the switching equipment and the Frame Relay network designers.

Figure 1-5 Frame Relay Switch Response to Congestion



Cisco routers respond to FECN by returning a test frame without any data, but with the BECN bit set. This relays the congestion condition to the source of the congestion. However, if the router detects a packet ready for transmission, it will set the BECN bit on it instead of using a test frame.

Upon receiving a BECN bit, Cisco routers respond in one of two ways. If the router is using DLSw+ Lite, then the LLC2 send window on the Frame Relay WAN is divided by two. Halving the send window causes the P/F bit in the LLC2 header to be set every *window*/2 frames. For example, if the window size is seven, then the P/F bit will be set every three frames instead of seven frames. Setting the P/F bit forces the remote link station in the Data Terminal Equipment (DTE) device, the remote receiving router, to acknowledge receipt of data more frequently, which ultimately slows down the sending router as it waits for acknowledgments. During periods of congestion this is important because it slows down the rate of transmission and also verifies receipt more frequently. Each time a BECN is received, the window size is halved until it reaches one. If eight frames are received without BECN, the window size is increased by one until the maximum window size is reached again.

The second way that a Cisco router responds to BECN requires Frame Relay traffic shaping. When traffic shaping is enabled and a BECN frame is received, the rate of transmission is multiplied by 75 percent or reduced by 25 percent. This occurs once every time interval (Tc) until a user-defined lower rate is reached (Tc = committed information rate [CIR] / access rate (AR)). If no BECN is received for eight consecutive intervals, the rate is increased by 1/8 until the maximum rate is reached (excessive or peak information rate [EIR or PIR]) or another BECN is detected. Thus, traffic rates fluctuate between and upper (EIR, PIR) and lower traffic rate (CIR or minimum information rate [MIR]) depending on the arrival rates of frames with BECN set on them.

Consolidated Link Layer Management (Router ForeSight)

BECN relies on the presence of backward (reverse direction) traffic to set the congestion notification bit. If there is no backward traffic, then no notifications can be sent. In addition, even with ample backward traffic, the timeliness of reverse traffic may not be adequate. With simplex traffic (for example, a video broadcast) the returning traffic on which to set the BECN bit will not be present often enough to provide congestion indications in a timely manner. To resolve this, an enhancement to the existing signaling protocol was made. This signaling enhancement, called consolidated link-layer management (CLLM), enables the Frame Relay network to more efficiently relay congestion notifications to the DTE (or router).

ForeSight is an advanced traffic management feature that reacts quickly to the slightest hint of congestion. On Frame Relay backbones, ForeSight detects congestion on a trunk, egress port, or virtual circuit and adjusts admission rates in the network at the point of ingress for a virtual circuit. Using CLLM, ForeSight can be extended to the router. Cisco calls this router feature Router ForeSight.

Cisco WAN switches send a CLLM congestion notification message every 100 ms on the reserved data-link connection identifier (DLCI) 1007. This notification rate permits ForeSight to efficiently update the router about congestion in the Frame Relay network. The switches send a single CLLM message to the Cisco router with a congestion state indication (a single bit for each DLCI) and cause of the congested state for each DLCI. If more DLCIs exist than can fit in a single CLLM message, then multiple messages are sent. For more information about CLLM, refer to ANSI T1.618 or ITU Q.922 Annex A.

Implicit Congestion Notification

Implicit congestion notification occurs when a frame is dropped. A silently discarded frame and a measurable increase in network delay are clear indications of network congestion to TCP/IP. Before viable explicit congestion notification schemes were implemented, TCP/IP existed on Frame Relay backbones because of the slow-start and back-off algorithms it implements. TCP/IP reacts to varying round trip times (RTT) and packet drops by adjusting the send window for a session. With slow-start algorithms, TCP/IP will eventually reach the rates (using the sliding window algorithm) that the Frame Relay virtual circuit can sustain. During periods of congestion, when a dropped packet is detected or RTT increases TCP reduces its transmission rate by reducing its send window. Through a series of computations involving RTT delays, sliding windows, and packet retransmission, TCP/IP will eventually reach a rate of transmission that approximates the CIR. Stateless nonwindowed protocols, such as UDP, and adverse network conditions, such as broadcast storms, do not fare well in a Frame Relay environment because there is no back-off algorithm.

CHAPTER 2

Maintaining SNA Response Times

Maintaining consistent and low SNA response times is the "holy grail" of the SNA Internetworking quest. The requirement is frequently called meeting service-level agreements or protecting mission-critical data. The origins of this requirement are often technical, historical, political, or some combination of the three. SNA response time requirements vary greatly between customers. This requirement does not limit itself to SNA applications. A plan for network migration to Frame Relay often requires equal attention be paid to protecting other mission-critical traffic such as a stock trading system. However, the techniques discussed in this guide are flexible and powerful enough to be adapted to the vast majority of customer settings, applications, and protocols.

The bulk of existing SNA networks are traditional hierarchical networks in which many remote sites access a single central site. Classic hub-and-spoke or star Frame Relay network topologies are common. Other design topologies, including variations on fully meshed or partially meshed networks, exist but are rare in comparison.

Latest industry studies show that approximately 95 percent of all public Frame Relay networks are star topologies, and we believe that this is also true for enterprise networks. The majority of line speeds are 64 kbps or lower. This is likely to create a common set of problems that are likely to occur as legacy SNA networks continue to migrate to the classic Frame Relay star topologies.

In this chapter, we will look at the most prevalent problems that degrade response times and techniques to resolve them. Much is written about the factors that contribute toward degrading response times. Most of the technical literature concentrates on transmission theory, error rates, retransmission algorithms, and queuing delays from a Layer 2 perspective. We will concentrate primarily on queuing delays in the router from a Layer 3 perspective. In this discussion, we will take the 80:20 approach to resolving problems and discuss the 20 percent of the issues that can resolve 80 percent of the response time problems.

Issues that Negatively Impact Response Times

This section discusses the issues that have a negative impact on response times.

Propagation Delay

Propagation delay is the amount of time it takes for a signal to propagate across the media on which it is transmitted. Propagation delay is fixed and subject to distance factors. A complete discussion of this is beyond the scope of this document. Coast-to-coast propagation delay typically accounts for approximately 65-75 msecs. SNA user response times are on the order of second intervals. Propagation delay is insignificant relative to the overall user perceived response times.

Transmission Delay

Transmission delay, or queueing delay, is the amount of time required to put data on the wire. It is sometimes referred to as serialization delay for a single packet. The major issue that impacts transmission time is the maximum transmission unit (MTU). For example, it takes almost .25 seconds to transmit 1500 bytes on a 64-kbps link. A queue build-up of ten or twenty 1500-byte packets would take two or three seconds to transmit. If an SNA packet was scheduled behind all this traffic, it would add at least that much time to the response times. Queuing delays during periods of congestion are a major factor that impacts the response times perceived by end users. The sections "Mixing Batch and Interactive Traffic" and "Mismatched Access Speeds" address resolving these delays.

Encapsulation overhead also contributes to transmission delays though not significantly. The issue of encapsulation overhead receives far too much attention. For example, DLSw+ TCP/IP encapsulation adds about 56 bytes of transmission overhead. It takes .007 seconds to transmit 56 bytes on a 64-kbps link. Not many users will notice this response time delay.

Latency

Latency is the delay associated with making a switching decision. CPU overhead is a contributing factor in latency. CPU overhead is the amount of time it takes the router to encapsulate and switch a packet (or frame). These factors contribute to latency in the Frame Relay backbone network on switches as well, but in smaller amounts.

Mixing Batch and Interactive Traffic

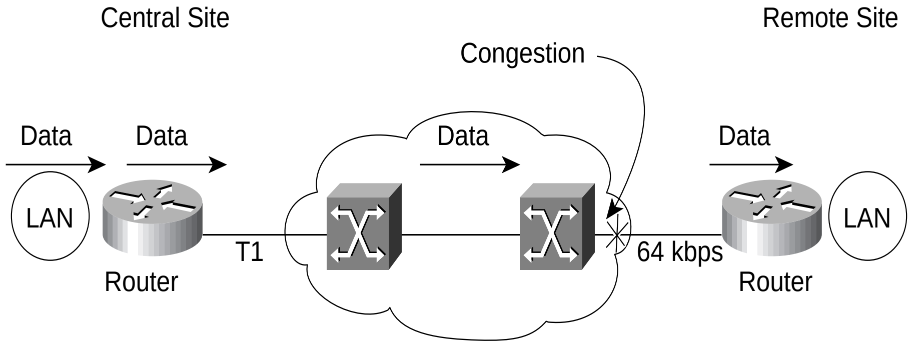
To preserve SNA response times, interactive SNA traffic must be separated from batch traffic and the interactive traffic must be prioritized. The initial reaction when addressing response time delays is to separate SNA from IP. For the most part, this would be correct. However, not all SNA traffic is interactive and not all batch traffic is IP (SNA can perform file transfers and IP has interactive applications). It is not the mix of protocols that causes the delays but the mix of batch and interactive traffic. Batch traffic interferes with interactive traffic response times.

To improve response times, you must separate batch traffic from interactive traffic and use a method of traffic prioritization to meet response time requirements. For example, if IND$FILE is present in a network, then even if SNA is high priority, the batch SNA traffic will negatively impact the response times for interactive SNA traffic.

Mismatched Access Speeds

Mismatched access speeds introduce a subtle problem that impacts just about all Frame Relay networks in production today. For example, consider traffic going from the central site to a remote site (see Figure 2-1). If a remote site is connected to a 64-kbps access speed line and the central site is connected to a T1 rate access speed line, traffic being transmitted by the router at the central site is clocked out of the Frame Relay serial port at T1 speeds. Thus, the likelihood of congestion appearing on the T1 interface is small. Because the data is traveling from T1 speeds to 64 kbps, a bottleneck will occur at some point. The bottleneck is encountered not at the T1 port when it exits the router, but at the 64-kbps egress port when it exits the Frame Relay switch. During periods of congestion at the Frame Relay switch egress port, frames are not prioritized and, because of excessive queue depths, transmission delays will be common. During periods of extreme congestion and buffer shortages, frames are randomly dropped at the egress port. If one of the randomly dropped frames is an SNA packet, this will cause a retransmit that will wreak havoc on response times.

Figure 2-1 Point of Congestion Caused by Mismatched Access Speeds



Techniques to Improve Response Times

The previous section talked about the issues that contribute to the degradation of response times. In this section, we talk about some generic techniques and some specific features of Cisco IOS software that can be used to improve SNA response times.

Reducing Transmission Delay

As previously mentioned, transmission delay is largely a constant based on the end-to-end propagation delay and the clock speed of the access port. The only way a user has to reduce transmission delay is to create smaller packets or increase clock rates. In many cases, when migrating an SNA network from a legacy SDLC 9.6-kbps or 19.2-kbps environment, significant improvements can be realized from the increase in clock speeds that Frame Relay provides.

Another alternative to lowering transmission delay is to reduce the IP MTU on the Frame Relay port. This improves response times by reducing the overall queuing delays. However, this can have a severe impact on the network by causing excessive IP fragmentation. Fragmenting not only causes CPU overhead to perform the fragmenting operation, but it also significantly increases the number of packets in transit on a enterprise network. Reducing the MTU from 1500 to 250 bytes effectively increases packet rates five-fold. This is not a major impact at remote sites where there are low rates of traffic, but at the central site aggregation points this can have a serious impact.

IP MTU path discovery solves the IP fragmentation problem, but it does not solve the problem of increased packet rates at central points of aggregation.

The solution to this dilemma is to fragment the frame just at the points where there are slow transmission speeds and to reassemble the fragments before forwarding the frame into the enterprise network, in other words a Layer 2 fragmenting scheme. The idea is to fragment the large frames, the so-called jumbo-grams, and interleave the shorter high-priority frames between the fragments. Dropping the MTU size is a Layer 3 fragmenting technique discussed above. There are two ways to fragment at Layer 2: FRF.12 and MLPPP. At this time, Cisco IOS software does not support FRF 12, but it will in a future release. Some Layer 2 multiservice Cisco products do support a variation of FRF 12 (for example, the Cisco 3810) to solve a similar problem for voice prioritization. At this time, the 3810 does not support any other Layer 3 QoS features (that is PQ, CQ, WFQ, and so forth). Cisco does support MLPPP in combination with weighted fair queueing (WFQ) and Resource Reservation Protocol (RSVP) on leased-line Point-to-Point Protocol (PPP) links, but not over Frame Relay (support for this is targeted in Release 12.0(4)T).

Another way to reduce the packet size is to compress the packet using Frame Relay payload compression or TCP header compression. Payload compression is a CPU-intensive process because it must compute mathematically complex operations on every bit of every byte transmitted. Keeping up with this demand at high packet rates is difficult. To make matters worse, the locations where compression is needed most are at remote sites where there are very slow lines and low-cost machines. Fortunately, the low-speed lines create a bottleneck that allows low-cost machines to keep up with the CPU processing demands. Care must be taken when considering using compression based on the platform in use. For example, a Cisco 2500 compressing traffic at 256-kbps rates runs at 99 percent CPU utilization. This situation must be avoided, so the recommendations on which speed to run on which platforms are as follows:

- Cisco 2500: less than 128 kbps
- Cisco 4000: less than 256 kbps
- Cisco 4500: less than 800 kbps
- Cisco 4700: less than T1 or E1 speeds
- Cisco 7x00: use compression service adapter (CSA)

Because the CPU cost is on a per-bit basis, the speeds listed are aggregate speeds. For example, on a Cisco 4700 one can compress on three 512-kbps links roughly totaling a single T1.

There are two forms of Frame Relay payload compression available. The oldest form is packet-by-packet compression, which is licensed from STAC Corporation. It implements a compression methodology that looks for repeated patterns on each packet. Thus, the larger the packet, the better the compression ratios as the likelihood of finding repeated characters increases.

The other form of payload compression is FRF 9, which is a standards-based compression method approved by the Frame Relay forum. FRF 9 is a dictionary-based compression methodology that uses a sliding window of the most recently transmitted data (the dictionary) to find repeated patterns. The dictionary is often much larger than a single packet and thus it can achieve much higher compression ratios (typically 2:1 ratios).

TCP header compression is another form of compression that is not as CPU intensive because it compresses only the 40-byte TCP/IP header. The benefits realized using TCP header compression are not of great significance if the packet is large because the ratio of payload to header is small. If packets are small, then header compression is much more beneficial. There is no known performance data available for TCP header compression.

Techniques for Separating Batch Traffic from Interactive Traffic

This section is not specific to Frame Relay environments. Many of the features and configuration samples discussed here are found in *DLSw+ Design and Implementation Guide*. This section provides additional information on design issues and, because DLSw+ is being implemented in large scale in Frame Relay environments, stresses the importance of Frame Relay.

To protect SNA response times, SNA traffic must be isolated from other ambient traffic present at a bottleneck. In general, batch traffic must be separated from interactive traffic. In some cases, this is as simple as separating SNA from IP, and will work in most circumstances. However, a more general solution is to separate all interactive traffic from the batch traffic. Interactive traffic can be in the form of SNA 3270, Telnet, Hypertext Transfer Protocol (HTTP), or other custom applications. Batch traffic can be either SNA 3270 (IND$FILE) or some form LU 6.2 or LU 0 file transfer application, FTP, NetBIOS, HTTP (yes, it can be considered batch or interactive), or any other custom application.

By classifying all interactive traffic together, bandwidth demands should be very meager in that traffic classification. Mixing interactive SNA 3270 with Telnet should not be a major issue because the demands of Telnet are minimal (mostly single-byte frames). There is also the option to create another class of service for Telnet if requirements call for it.

The more difficult task is to separate SNA batch traffic from SNA interactive traffic. Where dependent LUs are used the only means the router has to separate traffic is the SNA Network Addressable Unit (NAU), which is otherwise referred to as the LU address or the LOCADDR. Architecturally, SNA refers to this class of LUs as “dependent” because of their dependency on the VTAM SSCP. ALL LU type 2, all LU type 0, and some LU type 6.2 applications are dependent.

The separation of batch from interactive is further complicated if LU 6.2 parallel sessions is used because the NAU or LU address is dynamically allocated. Parallel session support is referred to as “independent” LUs because they do not require the assistance of the VTAM SSCP; it is APPN peer-to-peer communications. SNA class of service (CoS) to DLSw+ type of service (ToS), introduced in Cisco IOS Release 11.3, resolves this problem. This works by mapping SNA CoS to IP ToS (or precedence). SNA CoS to DLSw+ ToS can be used in conjunction with DLUR to migrate old dependent LUs to APPN. However, to map SNA CoS to DLSw+ ToS requires the use of APPN and DLSw+ on the same router and APPN in the enterprise network.

Features to separate SNA batch and interactive traffic without using APPN CoS to DLSw+ ToS exist in Cisco IOS. SNA batch traffic is typically a printer data stream or file transfer. Using the LU address, the router can redirect batch traffic from a particular LU (for example, LOCADDR 04) to a specific DLSw+ TCP port (1980) and redirect interactive traffic (LOCADDR 02) on a different DLSw+ TCP port (2065). Once the traffic is separated on different TCP flows, it is simply a matter of using any one of the various queuing techniques to establish packet scheduling priorities based on the DLSw+ TCP ports.

Special considerations must be made if IND$FILE is present in a network. It is important to highlight that the use of IND$FILE degrades SNA response times even in the legacy native SNA environments. IND$FILE is a very old, very inelegant method of transferring data. It essentially uses 3270 screens to batch move data. 3270 screens were not designed for this purpose. In networks today there is little excuse for using IND$FILE for data transfer. If IND$FILE is in use, there is no way to stop a user from starting a file transfer on any SNA 3270 session that could just as well be interactive. The best advice for customers trying to preserve SNA response times is to recommend they discontinue the use of IND$FILE and move to a more advanced method of file transfer. Even if a customer uses LU 6.2 for file transfers, SNA 3270 traffic can be identified in the dependent LU address space and be assigned to a DLSw+ TCP port. Then set the default for the remaining independent LU batch traffic to a different DLSw+ TCP port and prioritize at will.

Opening Multiple DLSw+ Ports

The first step in separating traffic on different flows is to create the flows on which to put the traffic. This is done by adding the priority keyword on the **dlsw remote-peer** command as follows:

```
dlsw remote-peer 0 tcp 1.1.1.1 priority
```

The priority keyword instructs DLSw+ to open four sessions to the remote peers on ports 2065, 1983, 1982, and 1981. By convention alone, these ports are given names: high, normal, medium, and low. This convention is not widely known, but it is critical to understanding how traffic is mapped to TCP port numbers. In reality, not much prioritization occurs until a queuing method is put in place on an output interface. The Cisco IOS software TCP driver checks port queues in descending numeric order. In fact, Cisco IOS software queuing methods can override this convention by mapping the high port (2065) to a low service class using priority queuing (there's really no need to do this in practice though). It is important to understand the naming conventions and the mapping of names to port numbers because it is not apparent from looking at the configuration commands.

This process will increase the number of aggregate ports that are open and active by a factor of four. Although the total number of TCP sessions is not considered an important factor when sizing for performance, the total count of TCP sessions at the central site needs to be watched. There are some scaling issues when session counts reach large numbers (600 sessions).

Mapping Traffic by SAP to DLSw+ Ports

On workstations at the LAN, instances of Layer 3 or higher protocol stacks access link layer (or data-link control) services through a SAP. The SAP serves two purposes, which can be confusing. It identifies the protocol and addresses the link. The IEEE 802 committee has administered a number space for SAPs that appear on LANs. For example, SNA has reserved SAP 0x04, 0x08, and 0x0C, and NetBIOS has SAP 0xF0. Unfortunately, the SAP addressing was too small (four bits source and destination) to accommodate all protocol types migrated from Ethernet. Thus, a special SAP (0xAA) was reserved to extend the protocol identification. SAP 0xAA indicates that another header appears (called the SNAP header for sub-network access point) to represent the Ethernet protocol type field. For example, IP uses the SNAP SAP (0xAA) and within the SNAP header specifies the familiar 0x0800 protocol ID.

If you need to separate traffic by SAP, then the SAP prioritization feature can be used. Note that SAP prioritization does not understand SNAP headers. Most protocols in the SNAP field are routable. It is most common to separate SNA from NetBIOS using SAP prioritization. To do this, you need to first build the following list:

```
sap-priority-list 1 high ssap 04 dsap 04
sap-priority-list 1 low ssap F0 dsap F0
```

These commands define a global sap-priority-list and map SNA traffic to DLSw+ port 2065 and NetBIOS traffic to DLSw+ port 1981. You need to know that the high and low keywords map to these DLSw+ ports.

Then the **sap priority-list** command must be applied to an input LAN interface or applied to a bridging domain. The input LAN interface can be a Token Ring interface. The bridging domain can be a DLSw+ bridge group for Ethernet.

```
interface tokenring 0
sap-priority  1
```

or

```
dlsw bridge-group 1 sap-priority 1
```

This process applies to any SAP that can be bridged. A packet with a SAP belonging to a routable protocol (such as SAP 0xAA) will be routed before it is evaluated by the **sap priority-list** command. As a result, this feature can be applied to other protocols and services identified by LLC SAP.

Mapping Traffic by LU to DLSw+ Ports

The feature to separate traffic by LU is called LOCADDR prioritization and uses the **locaddr priority-list** command. The configuration syntax for this command is as follows:

```
locaddr priority-list 1 02 high
locaddr priority-list 1 05 low
```

This configuration maps every LU addressed 02 to DLSw+ port 2065 and every LU addressed 05 to port 1981.

Next, the **locaddr priority-list** command must be applied to an input LAN interface or bridging domain and can be either Token Ring or a DLSw+ bridge group for Ethernet.

```
interface tokenring 0
locaddr-priority 1
```

or

```
dlsw bridge-group 1 locaddr-priority 1
```

Map SNA CoS to IP ToS

The process of mapping SNA CoS to IP ToS involves two steps. First, IP precedence is automatically established by virtue of using the priority keyword in the **dlsw remote-peer** command. Second, APPN CoS to IP ToS is automatically established by using APPN on the router and using DLSw+ to bridge SNA traffic on the same router.

In the example above, SNA or NetBIOS is being bridged using DLSw+ and the priority keyword is used and DLSw+ opens the four sessions on ports 1981, 1982, 1983, and 2065. In Release 11.3 and later, DLSw+ also assigns default IP precedent values to these TCP session ports, as shown in Table 2-1.

Table 2-1 TCP Port-to-IP Precedence Default Mapping

TPC Port	Priority Queue	IP Precedence	Precedence Numeric Value
2065	High	Critical	5
1981	Medium	Flash override	4
1982	Normal	Flash	3
1983	Low	Immediate	2

The default precedence values can be overridden using the **dlsw tos map** command or using policy-based routing. (See “Weighted Fair Queuing”.) The **dlsw tos map** command looks like the following:

```
dlsw tos map high 5 medium 2 normal 1 low 0
```

This example remaps medium priority traffic to immediate precedence, normal priority traffic to priority precedence, and low priority traffic to routine precedence. High-priority traffic remains at critical precedence.

The other method of mapping SNA CoS to IP ToS is when APPN is being routed and uses the DLSw+ VDLC interface as a link layer connection. This is not the same as bridging APPN over DLSw+. In this situation, the router is a full APPN Network Node routing SNA. The APPN Network Node participates fully in the session establishment process where it has easy access to session priority. When the Network Node accesses link layer services, the APPN CoS (APPN transmission priority) is mapped to a DLSw+ port. Table 2-2 lists the default mappings. There are conveniently four APPN transmission priorities and four DLSw+ priority ports as a result of the DLSw+ **priority** keyword.

Table 2-2 APPN CoS to IP ToS Mapping

APPN Mode Names	SNA Transmission Priority	TCP Port	Priority Queue	IP Precedence	Precedence Numeric Value
CPSNASVCMGR	Network	2065	High	Critical	5
#INTER	High	1981	Medium	Flash override	4

APPN Mode Names	SNA Transmission Priority	TCP Port	Priority Queue	IP Precedence	Precedence Numeric Value
#CONNECT	Medium	1982	Normal	Flash	3
#BATCH	Low	1983	Low	Immediate	2

Currently, there is no way to change these default mappings. For example, the mode CPSNASCVMG is assigned network transmission priority mapped to TCP port 2065, the mode #INTER is given high priority that is mapped to port 1981 at IP precedence flash override, #CONNECT is given medium priority, and #BATCH is given low priority. In addition BIND commands and IPM (pacing messages) are given their network transmission priority by APPN. See *APPN Design and Implementation Guide* and the *APPN Architecture Reference* (SC-3422) for more information on APPN transmission priorities and COS.

Choosing a Traffic Prioritization Method

After separating interactive and batch traffic into individual streams, a prioritization method needs to be chosen. Cisco IOS software has a large and growing list of prioritization options, including priority queuing (PQ), custom queuing (CQ), weighted fair queuing (WFQ), and weighted random early detection (WRED).

Prioritization methods collectively can be called “queuing methods,” “output queuing,” or “fancy queuing.” One important factor that is often overlooked is that there must be congestion on an output interface for traffic prioritization to take effect. By strict definition, “congestion” is when there are one or more packets on the interface output queue. Starting in the Release 11.0 time frame, these prioritization methods were fast-switched. At the first sign of congestion, when the hardware FIFO buffer on the serial chip is full or the CBUS transmit queue reaches a certain threshold, the fast switching software copies the packet into a system buffer and queues it on the output queue. The traffic waiting on this output queue gets the prioritization method applied to it. If the fast-switching software detects no congestion, then traffic is fast-switched.

Priority Queuing

This method of traffic prioritization is the oldest available. PQ has four classes of service: high, medium, normal, and low. Traffic is assigned to a service class using the **priority-list** command. An example of this command is as follows:

```
access-list 101 permit tcp any any eq 2065
access-list 101 permit tcp any eq 2065 any

priority-list 1 protocol ip high list 101
```

This example configures an extended access list to identify DLSw+ port 2065 traffic and puts it in the high PQ service class. All remaining traffic defaults to the normal service class queue. PQ will service each class queue in the order of priority in a round robin fashion until it is empty. The risk with using PQ is that the higher priority queues are given all bandwidth if it is required. If high-priority traffic is always present, no other service classes will get service and all other traffic is starved. This starvation issue historically has been the source of many network problems.

In the configuration example above, SNA traffic could potentially get all the bandwidth and preempt any other traffic from being transmitted. However, testing has shown that PQ can provide optimal response times (around .60 second response times on a 64-kbps link with 50 other file transfers active). PQ can be used effectively if the network designer is certain that SNA traffic will only be interactive with reasonable numbers of users sharing the link bandwidth. Ten users on a 64-kbps link is reasonable. 300 users on a 64-kbps link is not.

The final step in implementing PQ is to place the priority list on an outbound interface as follows:

```
interface serial0
priority-group 1
```

Custom Queuing

Because of the bandwidth starvation issues associated with using PQ, CQ was introduced to provide better transmission fairness and more service classes for sharing link bandwidth. Up to 16 service-class queues can be defined. The user can configure service levels for each queue by defining a byte count that limits the amount of bandwidth given to a service class. Just like PQ, traffic is designated for a particular service class by protocol or by using an access lists as follows:

```
access-list 101 permit tcp any any eq 2065
access-list 101 permit tcp any eq 2065 any
queue-list 1 protocol ip 1 list 101
queue-list 1 default 2
queue-list 1 queue 1 byte-count 2000
queue-list 1 queue 2 byte-count 8000
```

In this configuration, we have defined an access list to identify DLSw+ port 2065 traffic and place it in queue number 1, which has a byte count of 2000. All other traffic (including IP) will go to default queue number 2, which has a byte count of 8000. In this example, bandwidth is allocated in an SNA:IP ratio of 2:8, or 20 percent SNA and 80 percent IP (or other).

Using a byte count and describing CQ in terms of bandwidth allocation is unusual. CQ does not fragment packets, so it cannot stop transmitting in the middle of one. Thus, when defining queue byte counts, you must consider the MTU of the packets.

In the example above, if SNA has a maximum frame size of 1024 and IP the default MTU of 1500, then CQ would send two SNA frames followed by six IP frames. This achieves the desired bandwidth allocation. However, it also highlights one of the drawbacks of CQ, which is that CQ creates packet trains within the class and does not interleave packets well. The amount of time it takes to transmit six 1500-byte IP frames on a 64-kbps link is approximately 1.2 seconds. If the requirement were to maintain a subsecond response time, it would not be achieved.

If optimal response times are desired, it may be best to define byte counts in a way that the maximum size packet train is a single packet. This configuration is shown in the following example:

```
queue-list 1 queue 1 byte-count 1000
queue-list 1 queue 2 byte-count 1000
```

This configuration results in good packet interleaving but potentially gives SNA 50 percent of the bandwidth. In most circumstances SNA will never reach 50 percent bandwidth utilization because of its interactive nature. If SNA file transfers are present, they must be separated into their own queue. An SNA file transfer will demand the 50 percent utilization creating problems for interactive SNA users as well.

In summary, to achieve optimal response times, there is a risk/reward trade off. Using PQ you risk sacrificing the entire bandwidth (100 percent) to priority traffic. With CQ you risk losing half (50 percent) of the bandwidth.

The final step in configuring CQ is to place the CQ list on a serial interface as follows:

```
interface serial0
custom-queue-list 1
```

Weighted Fair Queuing

WFQ is a sophisticated queuing method. It is often referred to as a flow-based, as opposed to class-based, queuing method because it queues traffic per session (TCP or UDP port, and so on). Because of flow-based queuing, scaling WFQ on broadband trunks (speeds greater than E1) may be too CPU intensive. The distributed versatile interface processor (VIP) implementations are more effective in broadband trunk or LAN interface scenarios. Recently, distributed VIP implementations have introduced some class-based fair-queuing methods. However, these are not discussed in this document.

WFQ dynamically determines flow behavior in real-time and favors traffic with an interactive nature over other more aggressive traffic such as batch transfers. At first glance, this looks promising for SNA. However, although DLSw+ represents many interactive flows from many users, it is considered a single flow by WFQ. The concern is that many interactive SNA users will make the DLSw+ session appear bandwidth hungry and WFQ will penalize it. This is considered a major issue with using WFQ and DLSw+, but it is only a minor issue. Recall the interactive nature of SNA 3270 users. Even with 10 or 20 users over the same DLSw+ session, transaction rates are on the order of transactions per minute, which will result in packet rates of the same order. Under normal circumstances, this transaction rate will never reach levels that one would consider high bandwidth.

A greater threat to SNA response times has to do with TCP congestion management (slow-start and back-off). TCP congestion management uses a windowed protocol that automatically adjusts to the bottlenecks along the session path. In extreme circumstances (for example, sharing a 64-kbps link with 20 or 30 FTPs) TCP windows can be reduced to sizes that require every TCP packet to be acknowledged even during large file transfers (in other words, a window size of one). This situation very closely resembles an interactive session that has the adverse effect of creating artificial competition with interactive SNA DLSw+ traffic. Under these conditions WFQ cannot distinguish between the SNA interactive traffic and the TCP/IP batch traffic. As a result, administrative intervention is required to give SNA traffic differentiated service.

WFQ needs some method to distinguish SNA interactive traffic (encapsulated in TCP/IP) from other IP traffic during periods of extreme congestion. One method that can be used is to modify the weight of the DLSw+ traffic by using the precedence bits in the IP header. When WFQ determines the packet scheduling order, the lower the weight, the higher the packet priority. The computed weight is a function of the frame length (or transmission time of completion) and its position in its conversation queue. The length of the packet is reduced using the precedence bits in the IP packet ToS field. Thus, the only advantage SNA can hope to have over other traffic lies in gaining priority (lower weight) from the precedence bits. There are several mechanisms used to set precedence on SNA traffic.

Setting precedence on DLSw+ packets is done by using policy-based routing or by relying on DLSw+. (See “Map SNA CoS to IP ToS.”) Setting precedence on DLSw+ packets using policy-based routing can be done in the following manner:

```
ip local policy route-map SNADLSW

access-list 101 permit tcp any any eq 2065
access-list 101 permit tcp any eq 2065 any

route-map snadlsw permit 10
match ip address 101
set ip precedence critical
```

WFQ is enabled on a physical interface by default. No configuration is required.

Note: WFQ must be used in combination with traffic shaping over Frame Relay to be most effective.

Weighted Random Early Detection

WRED is one of the newer and more sophisticated queuing methods. It is considered a congestion avoidance method because it drops traffic based on mean queue depths instead of tail dropping as do the previous queuing methods. WRED is also considered a class-based queuing method because it deals with traffic based on class definitions. There are nine WRED classes. There is a class for each precedence level plus one for RSVP traffic. Currently, WRED works only for IP traffic. Traffic is placed on a single queue for transmission. Packets are selected for discard based on probabilities (thus the term random). The probability computations used for packet selection are a function of the packet precedence and mean queue depth on the output interface. The probability of a packet discard increases as the precedence decreases and as the mean queue depth increases.

By randomly selecting packets for discard based on probabilities, instead of tail-dropping when queues overflow, WRED resolves a phenomenon called global synchronization and more fairly discards among sessions using the link. Global synchronization can occur when simultaneously tail-dropping many packets across many sessions makes TCP back-off algorithms kick-in at the same time. TCP slow-start ramps up only to repeat the same pattern. This phenomenon has been observed on the Internet. Most enterprise networks today are not subject to this phenomenon, but it is possible in theory.

WRED works in conjunction with the precedence bits to differentiate service the same way WFQ does. Giving DLSw+ higher IP precedence reduces the probability that an SNA packet will be discarded, avoiding retransmissions that negatively impact response times. Note that SNA traffic is not scheduled at a higher priority using WRED. SNA is simply less likely to be dropped.

WRED is not recommended on slow links when using SNA because the single queue implementation can cause some queuing delays for SNA traffic and WRED does no traffic prioritization or packet sorting. WRED is more suitable for broadband trunks because the high speeds provide better algorithmic scaling properties and lessen the problems associated with long queuing delays.

Use the **random-detect** command to enable WRED as follows:

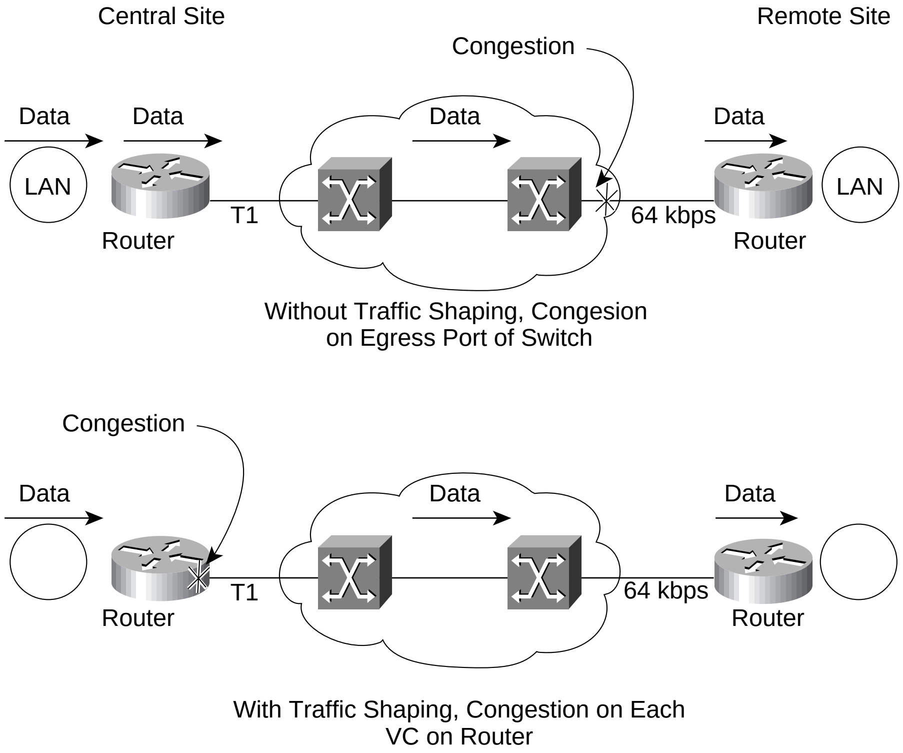
```
interface serial0
random-detect 9
```

Traffic Shaping

Frame Relay networks without traffic shaping cannot effectively prioritize traffic.

For an overview of congestion management in Frame Relay see the Technology Overview chapter. From that chapter, recall that mismatched access speeds can have a negative effect on SNA response times. Congestion occurs on the egress port of the Frame Relay switch where little traffic prioritization takes place. The solution to this problem is to use traffic shaping on the router and move the point of congestion to the router (see Figure 2-2). The concept is simple. Moving congestion from the Frame Relay switch to the router allows all of the above traffic prioritization methods to take place.

Figure 2-2 Effects of Traffic Shaping



An important clarification to make is that traffic shaping in a Frame Relay environment takes place on each DLCI in the router. So, there is a traffic-shaping output queue for every Frame Relay virtual circuit. Traffic shaping applied to each DLCI creates congestion (when traffic rates are higher than the traffic-shaping rates defined), and the congestion effectively creates an output queue for each DLCI. This per DLCI output queue is called the traffic-shaping queue. An output queuing method (such as PQ, CQ, WFQ, or the default FIFO) is applied to the traffic-shaping queue. Therefore, a traffic-shaping queue for each DLCI requires additional system buffer memory to accommodate the additional queuing. In situations where there are large numbers of DLCIs, buffer tuning may be necessary and additional I/O memory may be required on low-end systems that allocate buffer pool from I/O memory.

Another important issue to keep in mind when using traffic shaping in network designs is to ensure that traffic entering the Frame Relay network from the router, on the ingress port of the Frame Relay switch, is never faster than the traffic shaping taking place inside the Frame Relay network. For example, if a router is shaping at 70 kbps and the Frame Relay network is shaping at 64 kbps somewhere across the connection path inside the Frame Relay network there will be queuing delays. Also, it has been observed that Cisco IOS traffic shaping and ForeSight traffic shaping on Cisco WAN switches do not match. Even if both rates are set at 64 kbps, Cisco IOS software seems to send traffic faster than ForeSight. If an adjustment is not made for this difference, then an ingress queue buildup will result on the WAN switch running ForeSight. Packets will not be prioritized properly by the WAN switching equipment, so steps must be taken to ensure that the traffic rate entering the Frame Relay network is always lower than the rate made available by the Frame Relay network. This means traffic-shaping rates should generally be slightly lower on the router than the way they are defined on WAN switches.

There are two kinds of traffic shaping: generic traffic shaping (GTS) and Frame Relay traffic shaping (FRTS).

Generic Traffic Shaping

GTS can be applied only to a physical interface port or subinterface (in other words, internal to IOS an Interface Descriptor Block [IDB] style interface). In a Frame Relay environment, GTS should be used only if it is applied to a point-to-point subinterface. If there is a single DLCI present on each subinterface, it is equivalent to having traffic shaping applied to each DLCI. If the design calls for multipoint DLCIs, then GTS can be applied only to the IDB that holds them (on the major physical interface or subinterface), which is not very effective in a Frame Relay environment. GTS cannot accommodate per-DLCI traffic shaping when using multipoint (FRTS must be used instead).

GTS Traffic Prioritization

Only WFQ is supported on the GTS traffic-shaping queue. GTS also supports RSVP. There is no way to disable WFQ, so FIFO queuing is not supported. Regardless of limitations, GTS and WFQ may be the solution for some designs if the requirements are simple. Configure GTS as follows:

```
interface serial0.1 point-to-point
traffic-shape rate 62000 6000 6000 1000
traffic-shape fecn-adapt
frame-relay interface-dlci 801
```

In addition, GTS supports RSVP flows.

GTS Rate Adaptation

When GTS is applied to an interface with Frame Relay encapsulation, it will implement the rate adaptation algorithm discussed in the “Explicit Congestion Notification” section of the Technology Overview chapter. GTS also responds to FECN frames with a test frame. The commands to enable these features are as follows:

```
interface serial0.1 point-to-point
traffic-shape adaptive 32000
traffic-shape fecn-adapt
```

The adaptive traffic-shaping command will set the lowest rate that GTS will shape in the presence of severe congestion. The severity of the congestion is a function of the rate in which BECN frames are received.

Frame Relay Traffic Shaping

FRTS is also referred to as CIR rate enforcement. It can be used on single DLCI point-to-point subinterfaces or applied to specific DLCIs in a multipoint interface or subinterface. FRTS provides maximum flexibility by defining a map-class and configuring all the relevant traffic-shaping parameters for each DLCI in a multipoint interface or subinterface. The map-class can be applied to a group of multipoint DLCIs on an interface. There is also a provision for overriding a map-class for a particular DLCI via the **interface-dlci** command.

The following configuration shows how to apply a Frame Relay map-class to multipoint DLCIs and how to override the map on an individual DLCI for that interface (DLCI 401):

```
interface serial1/0
  encapsulation frame-relay ietf
  frame-relay traffic-shaping
  frame-relay class 16kcir
  frame-relay lmi-type ansi

interface serial1/0.1 multipoint
  frame-relay interface-dlci 401
  class 32kcir
  frame-relay interface-dlci 402
  frame-relay interface-dlci 403
  frame-relay interface-dlci 404
  frame-relay interface-dlci 405
  frame-relay interface-dlci 406
  frame-relay interface-dlci 407
  frame-relay interface-dlci 408
```

In the above configuration, DLCI 401 gets assigned the 32kcir map-class. All other DLCIs pick up the default map-class 16kcir from the major interface. The following configuration shows how to define the map-class statements:

```
map-class frame-relay 32kcir
  frame-relay cir 32000
  frame-relay bc 2000
  frame-relay be 2000
 map-class frame-relay 16kcir
  frame-relay cir 16000
  frame-relay bc 2000
  frame-relay be 2000
```

Frame Relay Traffic Shaping and Traffic Prioritization

FRTS also has a caveat with regard to traffic prioritization. Frame Relay traffic shaping supports FIFO, PQ, and CQ, but not WFQ. Cisco IOS software Release 12.0(4)T will support WFQ (but not RSVP). However, PQ, CQ, and FRTS in combination with multipoint DLCIs can be accomplished. In most large-scale Frame Relay implementations, FRTS is a requirement.

To enable PQ or CQ with FRTS, the PQ or CQ list is applied to the map-class. The following configuration shows how to apply PQ and CQ to a map-class:

```
map-class frame-relay 16kcir
  frame-relay cir 16000
  frame-relay bc 2000
  frame-relay be 2000
  frame-relay priority-group 1

 map-class frame-relay 16kcir
  frame-relay cir 16000
  frame-relay bc 2000
  frame-relay be 2000
  frame-relay custom-queue-list 1
```

Using traffic shaping and traffic prioritization methods on the routers is one way to prioritize traffic over a single Frame Relay virtual circuit or DLCI. It is practical because all traffic can be supported on single virtual circuit to each remote site. This is a low-cost solution. However, there are other methods to accomplish the same or similar prioritization objective. The next two sections discuss these alternate prioritization techniques.

Separating Traffic by DLCI

Traffic can be separated on different DLCIs using a feature called DLCI prioritization. This feature places the burden of providing QoS on the Frame Relay network instead of the router. The router does no prioritization of traffic. It is responsible for separating the traffic on different DLCIs. The Frame Relay switch does the prioritization based on DLCI. By doing this, users are relying on the Frame Relay QoS characteristics for the virtual circuit to meet service-level agreements and relying on the WAN switch to prioritize traffic. This feature in Cisco WAN switches is called priority PVC.

In many instances DLCI prioritization is a viable solution. The major deficiency with this solution is the extra cost of additional virtual circuits. In large Frame Relay networks, there is a scalability issue because of this as well. Also, the traffic prioritization mechanisms employed by Frame Relay switches are often crude. The switch creates two traffic priority service classes on the egress port of the Frame Relay switch: high and low. Cisco WAN switches using ForeSight can reduce the possibility of congestion, but when congestion does occur at the egress port, the best the switch can do is give the high-priority PVC a 10:1 ratio of service over the low-priority service class. For the most part, SNA will never need to use all 10 frames allocated to it on the high-priority virtual circuit. However, if severe congestion continues and frame buffers are depleted at the egress point, then SNA frames will be dropped with equal priority as the low-priority traffic.

To configure DLCI prioritization, the router reuses the same **priority-list** command as PQ. The difference is that the priority list is not placed on an outbound interface, but is placed in an interface command that determines which DLCI is designated as the high-priority DLCI, the medium-priority DLCI, the normal-priority DLCI, and the low-priority DLCI.

An example of DLCI prioritization is configured as follows:

```
interface serial 0
  ip address 1.1.1.1 255.255.255.0
  encapsulation frame-relay
  frame-relay priority-dlci-group 2 401 403 403 403

 access-list 101 permit tcp any any eq 2065
 access-list 101 permit tcp any eq 2065 any

 priority-list 2 protocol ip high list 101
```

In this example, DLCI 401 on the **frame-relay priority-dlci-group** command is given high-priority status. The IP extended access list, access-list 101, puts DLSw+ port 2065 traffic in the high-priority service class. Though not doing any actual prioritization, this configuration sequence establishes the relationship required to place DLSw+ port 2065 traffic into DLCI 401. DLCI 403 gets other traffic assigned to it and the Frame Relay switch must meet the QoS guarantees for the user.

Separating Traffic using Policy Routing

Policy-based routing (PBR) can also be used to direct the flow of traffic over Frame Relay. If Frame Relay point-to-point subinterfaces are in use, then PBR can be used to set the next interface to route the packet. Though no specific testing was done to verify, it is feasible to use policy routing to direct traffic over multipoint Frame Relay DLCIs by setting the next-hop address to the IP address in the Frame Relay map. The idea is the same as DLCI prioritization—direct incoming traffic to different DLCIs so the Frame Relay network can handle the QoS guarantees defined for each virtual circuit.

PBR is the preferred method to separate traffic on different DLCIs. PRB is fast switched (in Release 11.3 and later) and is easier to follow. There are some performance issues with PBR so care must be taken if it is to be used at high speeds. PBR will eventually be integrated with Cisco Express Forwarding (CEF) to address performance concerns. A sample configuration for PBR is as follows:

```
ip local policy route-map snadlsw

 interface serial1/0
  encapsulation frame-relay ietf
  frame-relay lmi-type ansi

 interface serial1/0.1 point-to-point
  ip address 1.1.1.1 255.255.255.0
  frame-relay interface-dlci 401

 interface serial1/0.2 point-to-point
  ip address 1.1.2.1 255.255.255.0
  frame-relay interface-dlci 402

 access-list 101 permit tcp any any eq 2065
 access-list 101 permit tcp any eq 2065 any

 route-map snadlsw permit 10
  match ip address 101
  set next-interface serial1/0.1

 route-map snadlsw permit 20
  set default next-interface serial1/0.2
```

In this example, DLSw+ port 2065 traffic goes out on subinterface serial1/0.1, which is assigned DLCI 401. All other traffic by default goes out subinterface serial1/0.2, which is assigned DLCI number 402.

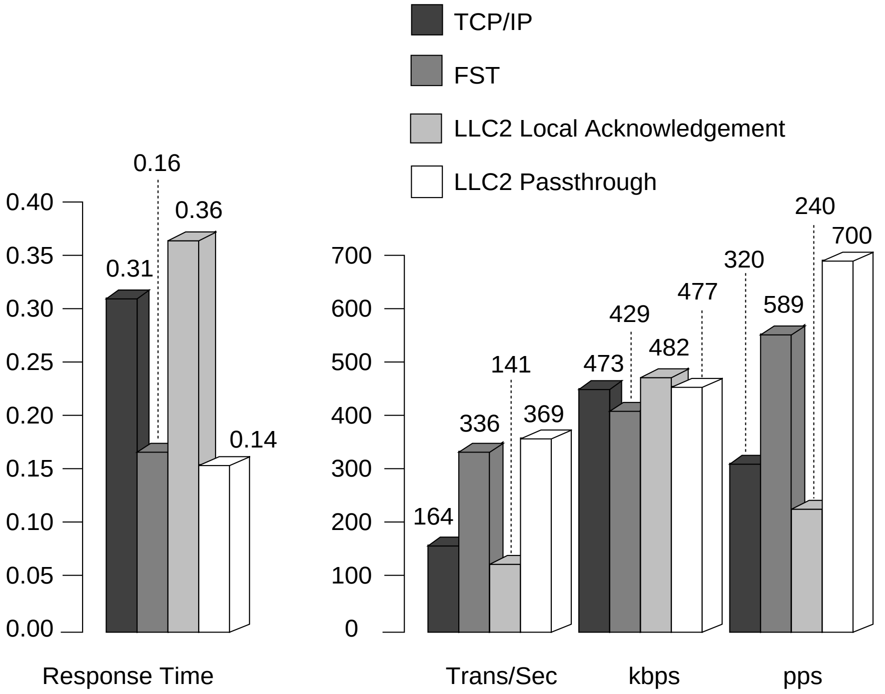
CHAPTER 3

Router Performance Considerations

This chapter discusses the relative performance characteristics of each DLSw+ encapsulation method available for Frame Relay. There is also a discussion about performance differences among the various encapsulation methods and how these differences might affect the decision-making process for network design.

Figure 3-1 compares the performance attributes of each encapsulation method for Frame Relay: TCP/IP, FST, LLC2 passthrough, and LLC2 LACK.

Figure 3-1 Relative Performance of DLSw+ Encapsulation Methods



These charts compare response times, transactions per second (Trans/sec), kbps throughput, and raw packets per second (pps) throughput. The response times are expressed in seconds using a transaction defined as 100 bytes in and 100 bytes out. Packet sizes have no bearing on transaction rates or packet throughput. It is a constant per-packet charge regardless of size. Response times and transaction rates reported are derived from the same transaction profile. The packets-per-second reported were derived from a test that sent 100-byte SNA packets as fast as possible from a PC. The kbps throughput reported was derived by simulating an SNA file transfer with 1024 maximum frame sizes. All the tests were run using a CIR of 512 kbps on a single PVC using a pair of Stratacom IPXs as Frame Relay switches.

General Guidelines

At remote sites, router performance is rarely an issue. At remote sites the processing power required to fill up low-speed lines is minimal. The remote line, not the CPU in the router, is the bottleneck.

However, at the central site, where there is typically a concentration of multiple high-speed lines (T1 or greater) handling traffic from a large number of remote sites, router performance is an issue. A considerable amount of traffic can flow through central site routers. The processor requirement for SNA features (like DLSw) is high, so it is important to select the correct router for the function it is to perform. It is useful to know how the various encapsulation methods impact performance on the router.

In general, when considering performance, an encapsulation method will fall into one of two major categories: LACK and passthrough. There are notable performance differences between the two. The functional benefits of running LACK are well known and are covered in detail in *DLSw+ Design and Implementation Guide*. However, the CPU overhead to reap these benefits comes at a cost. LACK terminates two connection-oriented sessions that provide guaranteed delivery service and data sequencing: the LLC session on the LAN and the TCP/IP or LLC session on the WAN. Software overhead to maintain these two sessions is not trivial and results in higher CPU utilization.

In contrast, passthrough modes of encapsulation have only the responsibility to encapsulate the packet and send it. Passthrough modes do not incur the additional CPU overhead of maintaining sessions. The CPU requirements to perform passthrough functions are far less than LACK. In addition, these differences in processor overhead impact the latency introduced by the two encapsulation styles. LACK modes have a higher latency than passthrough modes. Passthrough modes have better response times and higher transaction rates.

The first chart in Figure 3-1 shows how response times for the various encapsulation methods compare to each other. It shows that regardless of the specific encapsulation method, LACK and passthrough display similar properties.

The first chart also highlights another important point with respect to user-perceived response times. The relative differences in response time between LACK and passthrough are significant (roughly three-fold). However, the absolute differences are only on the order of .2 seconds. This amount of time is hardly perceptible to end users. Even in the worst case scenario, with routers running at 100 percent utilization, the difference in response time between LACK and passthrough modes is inconsequential to users. Queuing delays over slow links and congestion impact user-perceived response times.

The transaction rates and packet throughputs in the second chart in Figure 3-1 further demonstrates the impact that latency has on router performance. LACK and passthrough methods once again exhibit similar behavior. Passthrough is capable of significantly greater packet rates due to smaller processing overhead for each packet.

The second chart in Figure 3-1 also shows that file transfer rates are roughly equivalent for all encapsulation methods. This is because the packets are larger in a file transfer so the packet rates are lower. Latency is a per-packet cost so the lower packet rates have a lower total impact on performance. There is also a pipelining effect at the router. While a large packet is getting clocked out of the Frame Relay serial interface, the latency incurred for processing the packet happens in tandem. File transfer rates are affected more by encapsulation overhead on the wire than by latency. This effect is shown in the test results because LLC2 LACK encapsulation has the least overhead and the highest throughput. However, the differences are small and are well within the margin of error for the test. For all practical purposes, any encapsulation method can saturate the line bandwidth (in this test saturation was 512 kbps).

From the previous discussion we can conclude that passthrough modes of operation have about three times the packet-rate capacity of their LACK counterparts. In instances where bandwidth is abundant and CPU scarce, passthrough may be a viable option. However, in most circumstances CPU and performance is not an issue. Tests have shown that a Cisco 4700 router using TCP/IP encapsulation has the capacity to handle 1600 frames per second. This roughly translates to 400 transactions per second. Excluding any batch traffic, if users initiate a single transaction every minute, then a single Cisco 4700 is capable of handing 24,000 users. Modify the modeling parameters to one transaction every two minutes and the number of users double to 48,000. At one transaction every 10 minutes (a feasible scenario), the number raises an order of magnitude to 240,000 users.

To summarize the previous discussion:

- Requirements imposed by interactive users are modest because the traffic rate is limited by human interaction;
- The slowest most CPU-intensive encapsulation method can accommodate large numbers of interactive users;
- SNA batch traffic increases rates that greatly impact the number of users that can be supported.

Glossary

acknowledgment—Notification sent from one network device to another to acknowledge that some event (for example, receipt of a message) occurred. Sometimes abbreviated ACK.

all-routes explorer packet—Explorer packet that traverses an entire SRB network, following all possible paths to a specific destination. Sometimes called all-rings explorer packet.

AppleTalk—Series of communications protocols designed by Apple Computer consisting of two phases. Phase 1, the earlier version, supports a single physical network that can have only one network number and be in one zone. Phase 2, supports multiple logical networks on a single physical network and allows networks to be in more than one zone.

APPN—(Advanced Peer-to-Peer Networking) Enhancement to the original IBM SNA architecture. APPN handles session establishment between peer nodes, dynamic transparent route calculation, and traffic prioritization for APPC traffic.

backoff—The (usually random) retransmission delay enforced by contentious MAC protocols after a network node with data to transmit determines that the physical medium is already in use.

BECN—(backward explicit congestion notification) Bit set by a Frame Relay network in frames traveling in the opposite direction of frames encountering a congested path. DTE receiving frames with the BECN bit set can request that higher-level protocols take flow control action as appropriate.

BNN—(boundary network node) In SNA terminology, a subarea node that provides boundary function support for adjacent peripheral nodes. This support includes sequencing, pacing, and address translation.

CIR—(committed information rate) Rate at which a Frame Relay network agrees to transfer information under normal conditions, averaged over a minimum increment of time. CIR, measured in bits per second, is one of the key negotiated tariff metrics.

cluster controller—In SNA, a programmable device that controls the input/output operations of attached devices. Typically, an IBM 3174 or 3274 device.

CoS—(class of service) An indication of how an upper-layer protocol requires a lower-layer protocol to treat its messages. In SNA subarea routing, CoS definitions are used by subarea nodes to determine the optimal route to establish a given session. A CoS definition comprises a virtual route number and a transmission priority field. Also called ToS.

data-link control layer—Layer 2 in the SNA architectural model. Responsible for the transmission of data over a particular physical link. Corresponds roughly to the data-link layer of the OSI model.

DLCI—(data-link connection identifier) Value that specifies a PVC or SVC in a Frame Relay network. In the basic Frame Relay specification, DLCIs are locally significant (connected devices might use different values to specify the same connection). In the LMI extended specification, DLCIs are globally significant (DLCIs specify individual end devices).

DLSw—(data-link switching) An interoperability standard, described in RFC 1434, that provides a method for forwarding SNA and NetBIOS traffic over TCP/IP networks using data-link layer switching and encapsulation. DLSw uses SSP instead of SRB, eliminating the major limitations of SRB, including hop-count limits, broadcast and unnecessary traffic, timeouts, lack of flow control, and lack of prioritization schemes.

DSAP—(destination service access point) SAP of the network node designated in the Destination field of a packet.

DTE—(data terminal equipment) Device at the user end of a user-network interface that serves as a data source, destination, or both. DTE connects to a data network through a DCE device (for example, a modem) and typically uses clocking signals generated by the DCE. DTE includes such devices as computers, protocol translators, and multiplexers.

E1—Wide-area digital transmission scheme used predominantly in Europe that carries data at a rate of 2.048 Mbps. E1 lines can be leased for private use from common carriers.

explorer frame—Frame sent out by a networked device in a SRB environment to determine the optimal route to another networked device.

FECN—(forward explicit congestion notification) Bit set by a Frame Relay network to inform DTE receiving the frame that congestion was experienced in the path from source to destination. DTE receiving frames with the FECN bit set can request that higher-level protocols take flow-control action as appropriate.

FEP—(front-end processor) A device or board that provides network interface capabilities for a networked device. In SNA, an FEP is typically an IBM 3745 device.

FRAD—(Frame Relay access device) Any network device that provides a connection between a LAN and a Frame Relay WAN.

fragmentation—Process of breaking a packet into smaller units when transmitting over a network medium that cannot support the original size of the packet.

Frame Relay—Industry-standard, switched data-link layer protocol that handles multiple virtual circuits using HDLC encapsulation between connected devices. Frame Relay is more efficient than X.25, the protocol for which it is generally considered a replacement.

FST—(Fast Sequenced Transport) Connectionless, sequenced transport protocol that runs on top of the IP protocol. SRB traffic is encapsulated inside of IP datagrams and is passed over an FST connection between two network devices (such as routers). Speeds up data delivery, reduces overhead, and improves the response time of SRB traffic.

HDLC—(High-Level Data Link Control) Bit-oriented synchronous data link layer protocol developed by ISO. Derived from SDLC, HDLC specifies a data encapsulation method on synchronous serial links using frame characters and checksums.

HPR—(High-Performance Routing) Second-generation routing algorithm for APPN. HPR provides a connectionless layer with nondisruptive routing of sessions around link failures, and a connection-oriented layer with end-to-end flow control, error control, and sequencing.

LACK—(local acknowledgment) Method whereby an intermediate network node, such as a router, responds to acknowledgments for a remote end host. Use of local acknowledgments reduces network overhead and, therefore, the risk of time-outs. Also known as local termination.

latency—Delay between the time a device requests access to a network and the time it is granted permission to transmit. Delay between the time a device receives a frame and the time that frame is forwarded out the destination port.

LLC—(Logical Link Control) Higher of the two data-link layer sublayers defined by the IEEE. The LLC sublayer handles error control, flow control, framing, and MAC-sublayer addressing. The most prevalent LLC protocol is IEEE 802.2, which includes both connectionless and connection-oriented variants.

LOCADDR—Local address.

local explorer packet—Packet generated by an end system in an SRB network to find a host connected to the local ring. If the local explorer packet fails to find a local host, the end system produces either a spanning explorer packet or an all-routes explorer packet.

LU—(logical unit) A type of addressable unit in an SNA network. The LU is the port through which the end user accesses both the SSCP provided services, and communicates with other LUs at other nodes.

MAC—(Media Access Control) Lower of the two sublayers of the data-link layer defined by the IEEE. The MAC sublayer handles access to shared media, such as whether token passing or contention will be used.

MTU—(maximum transmission unit) Maximum packet size, in bytes, that a particular interface can handle.

NAU—(network addressable unit) An SNA term to describe the three entities which are addressable in an SNA network: SSCP, PU, and LU.

NCP—(Network Control Program) In SNA, a program that routes and controls the flow of data between a communications controller (in which it resides) and other network resources.

NetBIOS—(Network Basic Input/Output System) API used by applications on an IBM LAN to request services from lower-level network processes. These services might include session establishment and termination, and information transfer.

OUI—(Organizational Unique Identifier) 3 octets assigned by the IEEE in a block of 48-bit LAN addresses.

PAD—(packet assembler/disassembler) Device used to connect simple devices (like character-mode terminals) that do not support the full functionality of a particular protocol to a network. PADs buffer data and assemble and disassemble packets sent to such end devices.

priority queuing—Routing feature in which frames in an interface output queue are prioritized based on various characteristics such as packet size and interface type.

propagation delay—Time required for data to travel over a network, from its source to its ultimate destination.

PU—(physical unit) A type of addressable unit in an SNA network. Each node in the network has a PU, which provides services to control the physical configuration and the communication system resources associated with the node, and also to collect maintenance and operational statistics.

PVC—(permanent virtual circuit) Virtual circuit that is permanently established. PVCs save bandwidth associated with circuit establishment and tear down in situations where certain virtual circuits must exist all the time.

QLLC—(Qualified Logical Link Control) Data-link layer protocol defined by IBM that allows SNA data to be transported across X.25 networks.

QoS—Quality of Service.

queuing delay—Amount of time that data must wait before it can be transmitted onto a statistically multiplexed physical circuit.

RIF—(Routing Information Field) Field in the IEEE 802.5 header that is used by a source-route bridge to determine through which Token Ring network segments a packet must transit. A RIF is made up of ring and bridge numbers as well as other information.

RII—(Routing Information Identifier) Bit used by SRT bridges to distinguish between frames that should be transparently bridged and frames that should be passed to the SRB module for handling.

RSRB—(remote source-route bridging) A method of encapsulating SRB traffic over WAN links.

RSVP—(Resource Reservation Protocol) Protocol that supports the reservation of resources across an IP network. Applications running on IP end systems can use RSVP to indicate to other nodes the nature (bandwidth, jitter, maximum burst, and so forth) of the packet streams they want to receive. RSVP depends on IPv6. Also known as Resource Reservation Setup Protocol.

SAP—(service access point) A logical address that allows a system to route data between a remote device and the appropriate communications support.

SDLC—(Synchronous Data Link Control) SNA data-link layer communications protocol. SDLC is a bit-oriented, full-duplex serial protocol that has spawned numerous similar protocols, including HDLC and LAPB.

SNA—(Systems Network Architecture) An architecture designed by IBM to provide a unified systems design for their communication network products.

SNAP—(Subnetwork Access Protocol) Internet protocol that operates between a network entity in the subnetwork and a network entity in the end system. SNAP specifies a standard method of encapsulating IP datagrams and ARP messages on IEEE networks. The SNAP entity in the end system makes use of the services of the subnetwork and performs three key functions: data transfer, connection management, and QoS selection.

SRB—(source-route bridging) Method of bridging originated by IBM and popular in Token Ring networks. In an SRB network, the entire route to a destination is predetermined, in real time, prior to the sending of data to the destination.

SSAP—(source service access point) SAP of the network node designated in the Source field of a packet.

SSP—(Switch-to-Switch Protocol) Protocol specified in the DLSw standard that routers use to establish DLSw connections, locate resources, forward data, and handle flow control and error recovery.

star topology—LAN topology in which end points on a network are connected to a common central switch by point-to-point links. A ring topology that is organized as a star implements a unidirectional closed-loop star, instead of point-to-point links.

T1—Digital WAN carrier facility. T1 transmits DS-1-formatted data at 1.544 Mbps through the telephone-switching network, using AMI or B8ZS coding.

TDM—(time-division multiplexing) Technique in which information from multiple channels can be allocated bandwidth on a single wire based on preassigned time slots. Bandwidth is allocated to each channel regardless of whether the station has data to transmit.

ToS—Type of service.

traffic shaping—Use of queues to limit surges that can congest a network. Data is buffered and then sent into the network in regulated amounts to ensure that the traffic will fit within the promised traffic envelope for the particular connection. Traffic shaping is used in ATM, Frame Relay, and other types of networks. Also known as metering, shaping, and smoothing.

unnumbered frames—HDLC frames used for various control and management purposes, including link startup and shutdown, and mode specification.

virtual circuit—Logical circuit created to ensure reliable communication between two network devices. A virtual circuit is defined by a VPI/VCI pair, and can be either permanent (PVC) or switched (SVC). Virtual circuits are used in Frame Relay and X.25.

VTAM—(Virtual Telecommunications Access Method) Set of programs that control communication between LUs. VTAM controls data transmission between channel-attached devices and performs routing functions.

WAN—(wide-area network) Data communications network that serves users across a broad geographic area and often uses transmission devices provided by common carriers. Frame Relay, SMDS, and X.25 are examples of WANs.

WFQ—(weighted fair queuing) Congestion management algorithm that identifies conversations (in the form of traffic streams), separates packets that belong to each conversation, and ensures that capacity is shared fairly between these individual conversations. WFQ is an automatic way of stabilizing network behavior during congestion and results in increased performance and reduced retransmission.

References

This appendix contains a list of helpful reference documents.

- *ANSI T1.602 ISDN—Data-Link Layer Signaling Specification for Application at the User-Network Interface*
- *ANSI T1.606 Frame Relaying Bearer Service—Architectural Framework and Service Description*
- *ANSI T1.617 DSS1—Signaling Specification for Frame Relay Bearer Service*
- *ANSI T1.618 DSS1—Core Aspects of Frame Protocol for Use with Frame Relay Bearer Service*
- *FRF 1.1—User to Network Interface (UNI)*
- *FRF 2.1—Frame Relay Network to Network Interface (NNI)*
- *FRF 3.1—Multiprotocol Encapsulation*
- *FRF 4—Switched Virtual Circuits*
- *FRF 6—Frame Relay Service Customer Network Management (MIB)*
- *I.370—Congestion Management of the ISDN Frame Relaying Bearer Services*
- *ISO/IEC TR 9577:1996 Information Technology—Protocol Identification in the Network Layer*
- *Q.922 ISDN Data-Link Layer Specification for Frame Mode Bearer Services. ISDN DSS1—Signaling Specifications for Frame Mode Switched and Permanent Virtual Connection Control and Status Monitoring*
- *RFC 1293—Inverse Address Resolution Protocol*
- *RFC 1315—Management Information Base for Frame Relay DTEs*
- *RFC 1490—Multiprotocol Encapsulation*
- *RFC 1604—Definitions of Managed Objects for Frame Relay Service*
- *Supplement ANSI T1.606a-1992—Congestion Management and Frame Size*