

Introduction

This chapter describes Data-Link Switching Plus (DLSw+) and how to use this manual to design and configure a DLSw+ network. It reviews the key components of the DLSw features and describes the extensions to the standard that are included in DLSw+. Finally, it recommends how to proceed with designing your network.

DLSw+ Defined

DLSw+ is a means of transporting Systems Network Architecture (SNA) and NetBIOS traffic over a campus or WAN. The end systems can attach to the network over Token Ring, Ethernet, Synchronous Data Link Control (SDLC), Qualified Logical Link Control (QLLC), or Fiber Distributed Data Interface (FDDI). See Appendix B, “DLSw+ Support Matrix,” for details. DLSw+ switches between diverse media and locally terminates the data links, keeping acknowledgments, keepalives, and polling off the WAN. Local termination of data links also eliminates data-link control timeouts that can occur during transient network congestion or when rerouting around failed links. Finally, DLSw+ provides a mechanism for dynamically searching a network for SNA or NetBIOS resources and includes caching algorithms that minimize broadcast traffic.

In this document, DLSw+ routers are referred to as peer routers, peers, or partners. The connection between two DLSw+ routers is referred to as a peer connection. A DLSw circuit is comprised of the data-link control connection between the originating end system and the originating router; the connection between the two routers (typically a TCP connection), and the data-link control connection between the target router and the target end system. A single peer connection can carry multiple circuits.

DLSw+ supports circuits between SNA physical units (PUs) or between NetBIOS clients and servers. The SNA PU connectivity supported is PU 2.0/2.1-to-PU 4 (attached via any supported data-link controls), PU 1-to-PU 4 (SDLC only), PU 4-to-PU 4, and PU 2.1-to-PU 2.1 (any supported data-link control). See Appendix B for details about DLSw+ connectivity.

Note: DLSw+ provides support for redundant paths between PU 4 devices in Cisco IOS® Release 12.0 and later with the DLSw+ RIF Passthru feature. See the “Advanced Configuration” chapter for more details. Prior to Release 12.0, PU 4-to-PU 4 connectivity supports only a single path between front-end processors (FEPs) because of an idiosyncrasy in how FEPs treat duplicate source-route bridged paths.

DLSw Version 1 Standard

The DLSw standard was defined at the Advanced Peer-to-Peer Networking (APPN) Implementers Workshop (AIW) in the DLSw-related interest group. The DLSw Version 1 standard is documented in an informational Request for Comments (RFC), RFC 1795. RFC 1795 makes obsolete RFC 1434, which described IBM's original 6611 implementation of DLSw.

The DLSw standard describes the Switch-to-Switch Protocol (SSP) used between routers (called data-link switches) to establish DLSw peer connections, locate resources, forward data, handle flow control, and perform error recovery. RFC 1795 requires that data-link connections are terminated at the peer routers—that is, the data-link connections are locally acknowledged and, in the case of Token Ring, the routing information field (RIF) ends at a virtual ring in the peering router.

By locally terminating data-link control connections, the DLSw standard eliminates the requirement for link-layer acknowledgments and keepalive messages to flow across the WAN. In addition, because link-layer frames are acknowledged locally, link-layer timeouts should not occur. It is the responsibility of the DLSw routers to multiplex the traffic of multiple data-link controls to the appropriate TCP pipe and transport the data reliably across an IP backbone.

Before any end-system communication can occur over DLSw, the following must take place:

- Establish peer connection
- Exchange capabilities
- Establish circuit

After circuits are established, the standard describes how to control the flow of data between peers.

Establish Peer Connections

Before two routers can switch SNA or NetBIOS traffic, they must establish two TCP connections between them. The standard allows one of these TCP connections to be dropped if it is not required. (Cisco routers will drop the extra TCP connection unless they are communicating with another vendor's router that requires two TCP connections.) The standard also allows additional TCP connections to be made to allow for different levels of priority.

Exchange Capabilities

After the TCP connections are established, the routers exchange their capabilities. Capabilities include the DLSw version number, initial pacing windows (receive window size), NetBIOS support, list of supported service access points (SAPs), and the number of TCP sessions supported. Media Access Control (MAC) address lists and NetBIOS name lists can also be exchanged at this time, and if desired, a DLSw partner can specify that it does not want to receive certain types of search frames. It is possible to configure the MAC addresses and NetBIOS names of all resources that will use DLSw and thereby avoid any broadcasts. After the capabilities exchange, the DLSw partners are ready to establish circuits between SNA or NetBIOS end systems.

Establish Circuit

Circuit establishment between a pair of end systems includes locating the target resource (based on its destination MAC address or NetBIOS name) and setting up data-link control connections between each end system and its data-link switch (local router). SNA and NetBIOS are handled differently. SNA devices on a LAN find other SNA



devices by sending an explorer frame (a TEST or an exchange identification [XID] frame) with the MAC address of the target SNA device. When a DLSw router receives an explorer frame, the router sends a CANUREACH frame to each of the DLSw partners. If one of its DLSw partners can reach the specified MAC address, the partner replies with an ICANREACH frame. The specific sequence includes a CANUREACH_ex (explorer) to find the resource and a CANUREACH_cs (circuit setup) that triggers the peering routers to establish a circuit.

At this point, the DLSw partners establish a *circuit* that consists of three connections: the two data-link control connections between each router and the locally attached SNA end system, and the TCP connection between the DLSw partners. This circuit is uniquely identified by the source and destination circuit IDs, which are carried in all steady state data frames in lieu of data-link control addresses such as MAC addresses. Each circuit ID is defined by the destination and source MAC addresses and the destination and source link SAPs. The circuit concept simplifies management and is important in error processing and cleanup. Multiple DLSw+ circuits can flow over the same DLSw+ peer connection. After the circuit is established, information frames can flow over the circuit.

NetBIOS circuit establishment is similar, but instead of forwarding a CANUREACH frame that specifies a MAC address, DLSw routers send a name query (NetBIOS NAME-QUERY) frame that specifies a NetBIOS name. Instead of an ICANREACH frame, there is a name recognized (NetBIOS NAME-RECOGNIZED) frame.

Cisco's DLSw+ implementation caches information learned as part of the explorer processing so that subsequent searches for the same resource do not result in the sending of additional explorer frames.

Flow Control

The DLSw standard describes adaptive pacing between DLSw routers but does not indicate how to map this to the native data-link control flow control on the edges. The DLSw standard specifies flow control on a per-circuit basis and calls for two independent, unidirectional circuit flow-control mechanisms. It is important to have flow control for data traffic and because UDP Unicast handles explorer traffic, it is exempt from the flow control described here. Flow control is handled by a windowing mechanism that can dynamically adapt to buffer availability and end-station flow-control mechanisms.

Note: Cisco's DLSw+ implementation also uses TCP transmit queue depth to handle flow control.

Windows can be incremented, decremented, halved, or reset to zero. The granted units (the number of units that the sender has permission to send) are incremented with a flow-control indication from the receiver (similar to classic SNA session-level pacing). Flow-control indicators can be one of the following types:

- *Repeat*—Increment granted units by the current window size
- *Increment*—Increment the window size by one and increment granted units by the new window size
- *Decrement*—Decrement window size by one and increment granted units by the new window size
- *Reset*—Decrease window to zero and set granted units to zero to stop all transmission in one direction until an increment flow-control indicator is sent
- *Half*—Cut the current window size in half and increment granted units by the new window size

Flow-control indicators and flow-control acknowledgments can be piggybacked on information frames or can be sent as independent flow-control messages, but reset indicators are always sent as independent messages.

DLSw Version 2 Standard

The Version 2 standard is documented in RFC 2166. It includes RFC 1795 and adds the following enhancements:

- IP multicast
- UDP unicast
- Enhanced peer-on-demand routing feature
- Expedited peer connection

Users implement DLSw+ Version 2 for scalability if they are using multivendor DLSw devices with an IP multicast network. DLSw Version 2 requires complex planning because it involves configuration changes across an IP network.

IP Multicast

Multicast service avoids duplication and excessive bandwidth of broadcast traffic because it replicates and propagates messages to its multicast members only as necessary. It reduces the amount of network overhead in the following ways:

- Avoids the need to maintain TCP SSP connections between two DLSw peers when no circuits are available
- Ensures that each broadcast results in only a single explorer over every link

DLSw Version 2 is for customers who run a multicast IP network and do not need the advantages of border peering.

UDP Unicast

DLSw Version 2 uses UDP unicast in response to an IP multicast. When address resolution packets (CANREACH_ex, NETBIOS_NQ_ex, NETBIOS_ANQ, and DATAFRAME) are sent to multiple destinations (IP multicast service), DLSw Version 2 sends the response frames (ICANREACH_ex and NAME_RECOGNIZED_ex) via UDP unicast.

Enhanced Peer-on-Demand Routing Feature

DLSw Version 2 establishes TCP connections only when necessary and the TCP connections are brought down when there are no circuits to a DLSw peer for a specified amount of time. This method, known as peer-on-demand routing, was recently introduced in DLSw Version 2, but has been implemented in Cisco DLSw+ border peer technology since Cisco IOS Release 10.3.

Expedited TCP Connection

DLSw Version 2 efficiently establishes TCP connections. Previously, DLSw created two unidirectional TCP connections and then disconnected one after the capabilities exchange took place. With DLSw Version 2, a single bidirectional TCP connection is established if the peer is brought up as a result of an IP multicast/UDP unicast information exchange.

DLSw+ Features

DLSw+ is Cisco's implementation of DLSw. It goes beyond the standard to include the advanced features of Cisco's current remote source-route bridging (RSRB) and provides additional functionality to increase the overall scalability of DLSw.



DLSw+ includes enhancements in the following areas:

- *Scalability*—Constructs IBM internetworks in a way that reduces the amount of broadcast traffic and therefore enhances their scalability.
- *Availability*—Dynamically finds alternate paths quickly and optionally load balances across multiple active peers, ports, and channel gateways.
- *Transport flexibility*—Offers higher-performance transport options when there is enough bandwidth to handle the traffic load without risk of timeouts; in addition, the option to use lower-overhead solutions when bandwidth is at a premium and nondisruptive rerouting is not required.
- *Modes of operation*—Dynamically detects the capabilities of the peer router and operates according to those capabilities.
- *Management*—Works with enhanced network management tools such as CiscoWorks Blue Maps, CiscoWorks SNA View, and CiscoWorks Blue Internetwork Status Monitor (ISM).

DLSw+ Improved Scalability

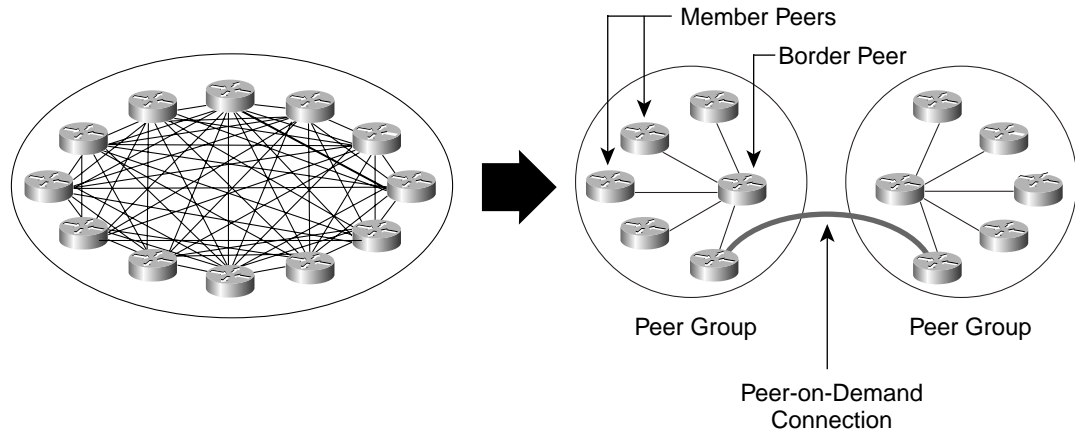
One of the most significant factors that limits the size of LAN internetworks is the amount of explorer traffic that traverses the WAN. There are several optimizations in DLSw+ to reduce the number of explorers.

Peer Group Concept

Perhaps the most significant optimization in DLSw+ is a feature known as *peer groups*. Peer groups are designed to address the broadcast replication that occurs in a fully meshed network. When any-to-any communication is required (for example, for NetBIOS or APPN environments), RSRB or standard DLSw implementations require peer connections between every pair of routers. This setup is not only difficult to configure, but it results in branch access routers having to replicate search requests for each peer connection. This wastes bandwidth and router cycles. A better concept is to group routers into clusters and designate a focal router to be responsible for broadcast replication. This capability is included in DLSw+.

With DLSw+, a cluster of routers in a region or a division of a company can be combined into a peer group. Within a peer group, one or more of the routers is designated to be the *border peer*. Instead of all routers peering to one another, each router within a group peers to the border peer and border peers establish peer connections with each other (see Figure 1-1). When a DLSw+ router receives a TEST frame or NetBIOS NAME-QUERY, it sends a single explorer frame to its border peer. The border peer checks its local, remote, and group cache for any reachability information before forwarding the explorer. If no match is found, the border peer forwards the explorer on behalf of the peer group member. If a match is found, the border peer sends the explorer to the appropriate peer or border peer. This setup eliminates duplicate explorers on the access links and minimizes the processing required in access routers.

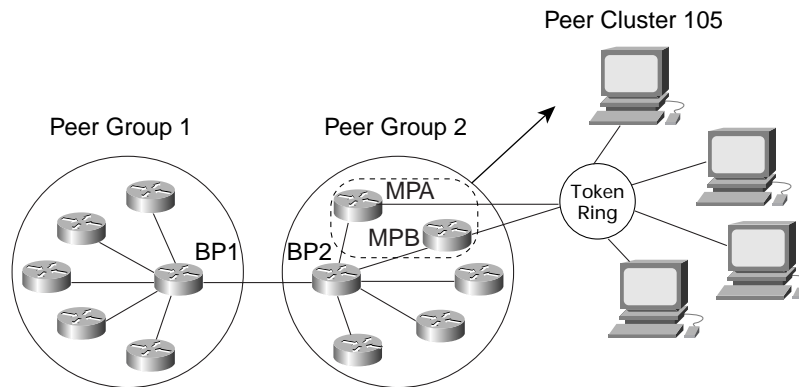
Figure 1-1 Peer Group Concept Simplifies and Scales Any-to-Any Networks



When the correct destination router is found, an end-to-end peer connection (TCP or IP) is established to carry end-system traffic. This connection remains active as long as there is end-system traffic on it, and it is dynamically torn down when not in use, permitting casual, any-to-any communication without the burden of specifying peer connections in advance. It also allows any-to-any routing in large internetworks where persistent TCP connections between every pair of routers would not be possible.

You can further segment routers within the same peer group that are serving the same LAN into a *peer cluster*. This segmentation reduces explorers because the border peer recognizes that it only has to forward an explorer to one member within a *peer cluster* (see Figure 1-2).

Figure 1-2 Peer Cluster Feature Reduces Explorer Traffic





Explorer Firewalls

To further reduce the amount of explorer traffic that enters the WAN, there are a number of filter and firewall techniques to terminate the explorer traffic at the DLSw+ router. A key feature is the explorer firewall.

An explorer firewall permits only a single explorer for a particular destination MAC address to be sent across the WAN. While an explorer is outstanding and awaiting a response from the destination, subsequent explorers for that MAC address are not propagated. When the explorer response is received at the originating DLSw+ router, all subsequent explorers are handled based on the newly acquired information. This eliminates the start-of-day explorer storm that many networks experience.

UDP Unicast

The UDP unicast feature eliminates unnecessary congestion caused by retransmission of explorers and unnumbered information frames because it sends those frames via UDP rather than TCP. Cisco's DLSw+ introduced the UDP unicast feature prior to the release of DLSw Version 2 in Cisco IOS Release 11.2(6)F. One difference between the two enhancements is that the Release 11.2(6)F UDP unicast feature requires that a TCP connection exist before packets are sent via UDP. Because the TCP session is up and capabilities are exchanged, the peers know exclusive reachability information that permits them to further reduce the explorer load on the network. DLSw Version 2 sends UDP/IP multicast and unicast before the TCP connection exists, which further propagates explorers.

NetBIOS Dial-on-Demand Routing

To further reduce the amount of traffic on a WAN, DLSw+ filters the NetBIOS session alive packets that are sent periodically over a WAN in a dial-on-demand routing (DDR) environment. These session alive packets do not require a response, do not impede proper data flow, and keep the DDR interfaces up. NetBIOS DDR reduces the number of unwanted per-packet charges that occur in a DDR network.

DLSw+ Enhanced Availability

One way DLSw+ offers enhanced availability is by maintaining a reachability cache of multiple paths for local and remote destination MAC addresses or NetBIOS names. For remote resources, the path specifies the peer to use to reach this resource. For local resources, the path specifies a port number. If there are multiple paths to reach a resource, the router will mark one path preferred and all other paths capable. If the preferred path is not available, the next available path is promoted to the new preferred path, and recovery over an alternate path is initiated immediately.

The way that multiple capable paths are handled with DLSw+ can be biased to meet the needs of the network:

- *Fault tolerance*—Biases circuit establishment over a preferred path, but also rapidly reconnects on an active alternate path if the preferred path is lost
- *Load balancing*—Distributes circuit establishment over multiple DLSw+ peers in the network or ports on the router

The default for DLSw+ is to use fault tolerant mode. In this mode, when a DLSw+ peer receives a TEST frame for a remote resource, it checks its cache. If it finds an entry and the entry is fresh (that is, if it is not verified within the last verify interval), the DLSw+ peer responds immediately to the TEST frame and does not send a CANUREACH frame across the network. If the cache entry is stale, then the originating DLSw+ peer sends a CANUREACH directly to each peer in the cache to validate the cache entries (this is known as a directed verify). The user configures the `sna-verify-interval` to determine the length of time a router waits before marking the cache

entry stale. See the “Customization” chapter. If any peer does not respond, it is deleted from the list. This may result in reordering the cache. The user configures the `sna-cache-timeout` interval to determine the amount of time that cache entries are maintained before they are deleted. See the “Customization” chapter.

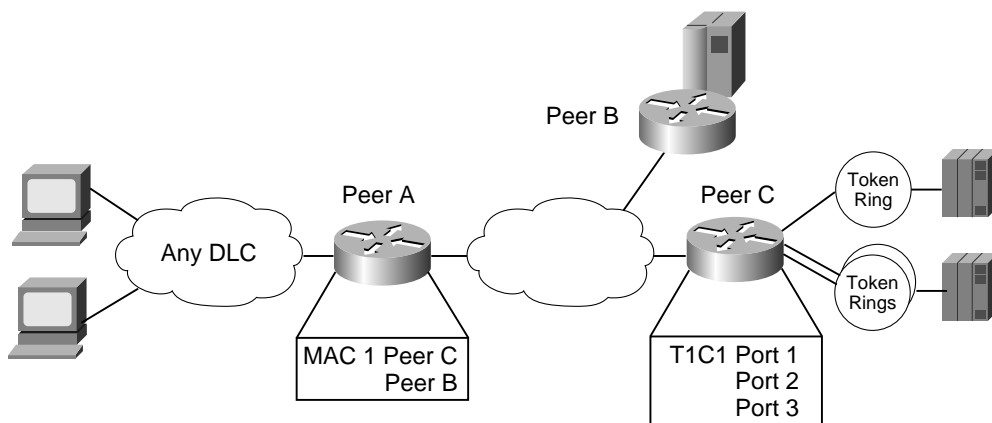
At the destination DLSw+ router, a slightly different procedure is followed using the local cache entries. If the cache entry is fresh, the response is sent immediately. If the cache entry is stale, a single route broadcast TEST frame is sent over the all ports in the cache. If a positive response is received, an ICANREACH frame is sent to the originating router. TEST frames are sent every 30 seconds. The user configures these timers. See the “Customization” chapter.

Alternately, when there are duplicate paths to the destination end system, you can configure load balancing. DLSw+ alternates new circuit requests in either a round-robin or *enhanced* load balancing fashion through the list of capable peers or ports. If round-robin is configured, the router distributes the new circuit in a round-robin fashion, basing its decision on which peer or port established the last circuit. If enhanced load balancing is configured, the router distributes new circuits based on existing loads and the desired ratio. It detects the path that is underloaded in comparison to the other capable peers and assigns new circuits to that path until the desired ratio is achieved.

This feature is especially attractive in SNA networks. A very common practice used in the hierarchical SNA environment is assigning the same MAC address to different mainframe channel gateways—for example, FEPs or Cisco routers with Channel Interface Processors (CIPs). If one channel gateway is unavailable, alternate channel gateways are dynamically located without any operator intervention. Duplicate MAC addressing also allows load balancing across multiple active channel gateways or Token Ring adapters.

DLSw+ ensures that duplicate MAC addresses are found, and it caches up to four DLSw peers or interface ports that can be used to find the MAC address. This technique can be used for fault tolerance and load balancing. When using this technique for fault tolerance, it facilitates a timely reconnection after circuit outages. When using this technique for load balancing, it improves overall SNA performance by spreading traffic across multiple active routers, Token Ring or FDDI adapters, or channel gateways, as shown in Figure 1-3. Load balancing not only enhances performance, it also speeds up recovery from the loss of any component in a path through the network because a smaller portion of the network is affected by the loss of any single component.

Figure 1-3 DLSw+ Caching Techniques Provide Load Balancing across Multiple Central Site Routers, Token Rings, and Channel Gateways



In addition to supporting multiple active peers, DLSw+ supports *backup peers*, which are only connected when the primary peer is unreachable.



DLSw+ Transport Flexibility

The transport connection between DLSw+ routers can vary according to the needs of the network and is not tied to TCP/IP. DLSw is tied to TCP/IP. Cisco supports five different transport protocols between DLSw+ routers:

- *TCP/IP*—Transports SNA and NetBIOS traffic across WANs where local acknowledgment is required to minimize unnecessary traffic and prevent data-link control timeouts and where nondisruptive rerouting around link failures is critical. This transport option is required when DLSw+ is operating in DLSw standard mode.
- *TCP/IP with RIF Passthru*—Transports SNA and NetBIOS traffic across WANs where the RIF is not terminated. This solution allows multiple active paths between FEPs.
- *Fast Sequence Transport (FST)/IP*—Transports SNA and NetBIOS traffic across WANs with an arbitrary topology; this solution allows rerouting around link failures, but recovery may be disruptive depending on the time required to find an alternate path. This option does not support local acknowledgment of frames.
- *Direct*—Transports SNA and NetBIOS traffic across a point-to-point High-Level Data Link Control (HDLC) or point-to-point Frame Relay (DLSw+ Lite) connection where the benefits of an arbitrary topology are not important and where nondisruptive rerouting around link failures is not required. This option does not support local acknowledgment of frames.
- *DLSw+ Lite*—Transports SNA and NetBIOS traffic across a point-to-point Frame Relay connection where local acknowledgment and reliable transport are important, but where nondisruptive rerouting around link failures is not required. DLSw Lite uses the RFC 1490 encapsulation of Logical Link Control, type 2 (LLC2). It is a form of Direct encapsulation.

DLSw+ Modes of Operation

Cisco has been shipping IBM internetworking products for many years. There is a substantial installed base of Cisco routers running RSRB today. Therefore, it is sometimes preferable for DLSw+ and RSRB to coexist in the same network and in the same router. In addition, because DLSw+ is based on the new DLSw standard, it must also interoperate with other vendors' implementations that are based upon that DLSw standard.

There are three different modes of operation for DLSw+:

- *Dual mode*—A Cisco router can communicate with some remote peers using RSRB and with others using DLSw+, providing a smooth migration path from RSRB to DLSw+. In dual mode, RSRB and DLSw+ coexist on the same box. The local peer must be configured for both RSRB and DLSw+; and the remote peers must be configured for either RSRB or DLSw, but not both.
- *Standards compliance mode*—DLSw+ can detect automatically (via the DLSw capabilities exchange) if the participating router is manufactured by another vendor; therefore operating in DLSw standard mode.
- *Enhanced mode*—DLSw+ can detect automatically that the participating router is another DLSw+ router, therefore operating in enhanced mode, making all of the features of DLSw+ available to the SNA and NetBIOS end systems.

Some of the enhanced DLSw+ features are also available when a Cisco router is operating in standards-compliance mode with another vendor's router. In particular, enhancements that are locally controlled options on a router can be accessed even though the remote router does not have DLSw+. These enhancements include load balancing, local caching (the ability to determine if a destination is on a LAN before sending CANUREACH frames across a WAN), explorer firewalls, and media conversion.

DLSw+ Management Features

DLSw+ supports several network management tools that enables network administrators to more easily troubleshoot and manage their network. For details on the supported applications, see the "Using CiscoWorks Blue: Maps, SNA View, and Internetwork Status Monitor" chapter.

How to Proceed

If you have a simple hierarchical network with a small volume of SNA traffic, read the “Getting Started” chapter, which describes what configuration commands are required in all DLSw+ implementations and provides configuration examples for SDLC, Token Ring, Ethernet, and QLLC. After reading the “Getting Started” chapter, you can read about advanced features, customization, and bandwidth management.

If you have a large hierarchical network (hundreds of branch offices), read the “Designing Hierarchical Networks” chapter, which will tell you how to determine the correct number and types of routers to place at the central site and discusses options for peer placement, peer backup, and broadcast reduction.

If you require any-to-any communication between NetBIOS or APPN applications, read the “Designing Meshed Networks” chapter, which describes border peer placement, numbers of peers per group, and how to minimize broadcast replication.

If you are starting with an RSRB network, read the “RSRB Migration and Multivendor Interoperability” chapter.

The “Using Show and Debug Commands” and “Using CiscoWorks Blue: Maps, SNA View, and Internetwork Status Monitor” chapters describe network management capabilities available with DLSw and should be read by all DLSw+ users.

The “DLSw+ Ethernet Redundancy Feature” chapter describe how to provide redundancy in an Ethernet environment. It describes how the feature works and how to configure it and provides sample configurations. If you are running an earlier release than 12.0(5)T, you should read Appendix C, “Ethernet Redundancy,” for a discussion on network design issues in a DLSw+ environment with Ethernet-attached end systems.

Finally, the “Using DLSw+ with Other Features” chapter describes how to use DLSw+ in conjunction with downstream physical unit (DSPU) concentration, LAN Network Manager, SNA Switching Services (SNASw), and native client interface architecture (NCIA).

The appendixes include memory requirements to assist in network planning and feature, media, and release matrices.