CHAPTER 9

Packet over SONET

SONET is a time-division multiplexing (TDM) architecture that was designed to carry voice traffic. All traffic in SONET is broken down into slots of 64-kbps DS0 increments. A DS0 is the voice line that is typically hard-wired into homes. TDM architectures are not ideal solutions for transporting data. Cable and DSL providers have shown this with their high-data-throughput broadband offerings that do not incur the same costs as comparable TDM services would incur.

When it was discovered that computer data could be transported over telephone circuits, service providers (SPs) leveraged their existing SONET rings. SONET rings were designed and deployed to transport voice but could transport voice by breaking down the data needs into manageable pieces and transporting in 64-kbps increments. When anything less than 100 percent of a TDM circuit was used, the remainder is stuffed with arbitrary data and therefore wasted from both the customer's and SP's perspective. Frame Relay technology offered statistical multiplexing, which offered a solution to the inefficiencies of SONET-based services. Unfortunately, the designers of Frame Relay did not have quality of service (QoS) in mind with the design of the technology. Many carriers also offered zero committed information rate (CIR) services only, which guaranteed the end user absolutely no class of service (CoS). Customers found this unacceptable.

ATM offered a solution to the QoS issues of Frame Relay and offered scalability in the optical carrier (OC-*n*) domain. ATM relied on a fixed-size cell that is not compatible with the Ethernet technologies that most LANs employ. ATM-designed hardware includes a segmentation and reassembly (SAR) layer to translate Layer 2 frames (Ethernet, Token Ring, FDDI, and so on) into ATM cells. The SAR functionality introduced slight delays in the network, but it is prohibitively expensive and complex to design. Because of the issues associated with ATM, many vendors have not deployed OC-192 ATM interfaces at this time. There is also a concept known as *cell tax* with ATM deployments. ATM introduces extra overhead into each transmission because of its fixed size of 53 bytes. If a Layer 2 (Ethernet) frame does not fall on a cell boundary, the rest of the cell is padded to meet the 53-byte cell requirement. ATM cells might be efficiently multiplexed into a SONET frame, but the architecture has delays and inefficiencies that must be accounted for.

Packet over SONET (PoS) is a highly scalable protocol that overcomes many of the inefficiencies of ATM, while providing legacy support to internetworks with existing SONET architectures. PoS provides a mechanism to carry packets directly within the SONET synchronous payload envelope (SPE) using a small amount of High-Level Data Link Control (HDLC) or PPP framing.

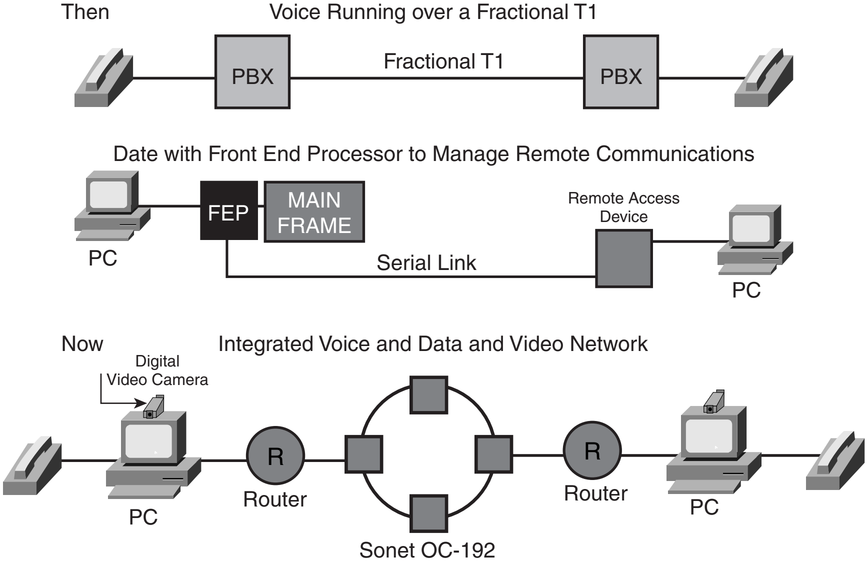
Evolution of Voice and Data Networks

Voice and data networking is constantly evolving as the technology evolves. After the telegraph, telecommunications networks evolved to transport the spoken word. The next evolutionary step, data networks, occurred in the mid-1900s with the advent of computers. Although data networking started out small because only the largest corporations could afford computers, computers have fallen to such a low entry-level price that most people can afford to have a computer now and to be connected to the Internet. Data networks have evolved to the point that the benefits of converging voice and data networks into the same data infrastructure can no longer be ignored. SPs that have legacy SONET infrastructures can still offer customers high-speed alternatives with technologies such as PoS.

The first communications systems were mainframe computers linked to dumb terminals. The Synchronous Data Link Control (SDLC) protocol, developed by IBM, made this system possible by allowing communication between a mainframe and a remote workstation over long distances. This protocol evolved into the High-Level Data Link Control (HDLC) protocol, which provided reliable communications and the ability to manage the flow of traffic between devices. HDLC is an open industry standard protocol, whereas SDLC is an IBM proprietary protocol that must be licensed by IBM. Industry standard protocols such as TCP/IP drive the adoption and low costs of telecommunications equipment.

Cisco offered an enhanced multiprotocol version of the HDLC protocol to enable various protocols over the HDLC (High-Level Data Link Control) Layer 2 framing. This Cisco HDLC protocol is proprietary and exclusive to Cisco. The HDLC standard was loose at the time of Cisco's creation and left too much room for interpretation. When this was standardized in RFC 1619 with PPP in HDLC-like framing, Cisco's HDLC protocol was not compliant. Point-to-Point Protocol (PPP) evolved from HDLC; it offers an industry-standard way to provide multiprotocol networking abilities, as well as many enhancements such as authentication, multilink, and compression. PPP is used for many other technologies, including ISDN. HDLC and PPP are scalable to architectures with fast speeds. Figure 9-1 shows this networking evolution.

Figure 9-1 *Voice and Data Network Evolution*


Then
Voice Running over a Fractional T1
PBX
Fractional T1
PBX
Date with Front End Processor to Manage Remote Communications
FEP
MAIN FRAME
Remote Access Device
PC
Serial Link
PC
Now
Integrated Voice and Data and Video Network
Digital Video Camera
R
R
PC
Router
Router
PC
Sonet OC-192

Applications for PoS

PoS is a Layer 2 technology that uses PPP in HDLC encapsulation, using SONET framing. The PoS solution lowers the cost per megabyte when compared to other Wide Area Networking architectures. The PoS interface supports SONET level alarm processing, performance monitoring, synchronization, and protection switching. This support enables PoS systems to seamlessly interoperate with existing SONET infrastructures and provides the capability to migrate to IP+Optical networks without the need for legacy SONET infrastructures. PoS is used in a point-to-point environment, much like the legacy T-carrier architectures, but without the need for TDM.

PoS efficiently encapsulates IP traffic with a low-overhead PPP header. When encapsulated, the traffic is placed inside an HDLC-delimited SONET SPE and transported across SONET. Voice, video, and data can be carried within the IP packets using Layer 3 QoS mechanisms to control priority when bandwidth contention occurs.

PoS can be used in tandem with other technologies carried over SONET architectures. PoS is not compatible with these other technologies, but is not aware of them because they are being transported over different time slots. PoS, TDM voice, ATM, and Dynamic Packet Transport (DPT) can each use their required synchronous transport signals, not interacting with each other. PoS interfaces are available in concatenated and nonconcatenated (channelized) options. Channelized interfaces are more costly than concatenated interfaces.

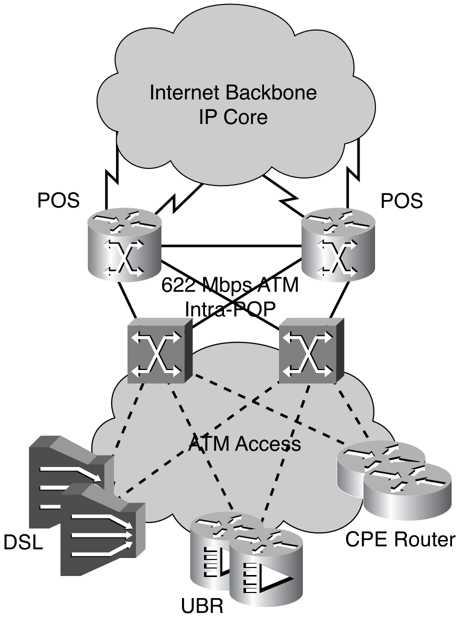
ATM and PoS

ATM and PoS can be used within the same network. ATM technology provides an effective, flexible provisioning mechanism for low- to high-speed network access. ATM switches can be used to aggregate digital subscriber line (DSL), cable, and customer traffic by using permanent virtual circuits (PVCs) or switched virtual circuits (SVCs), which can then feed into compatible downstream routers. This traffic is then fed into higher-speed links attached to the Cisco 12000 series router for transport through the core through PoS interfaces.

An advantage of ATM is its innate support for QoS. PoS is fully capable of supporting the transport of time-sensitive data, using Layer 3 mechanisms. Technologies such as Resource Reservation Protocol (RSVP), committed access rate (CAR), and Weighted Random Early Detection (WRED) enable providers to offer QoS solutions in a more cost-effective manner than ATM. These technologies are Layer 3 implementations for QoS. This book does not focus on QoS, but other Cisco Press books covering QoS are available (such as *IP Quality of Service*; ISBN: 1-57870-116-3).

Figure 9-2 is an ATM aggregation design in which switches are placed on the ATM network edge and translate ATM traffic into PoS traffic. Notice the amount of PoS interfaces required to create a resilient network design with no single point of failure. PoS is a point-to-point technology, regardless of the distance traveled. A point of presence (POP) environment where equipment might be closely located would be perfect for Very Short Reach (VSR) optics using PoS technology. VSR optics are lower in price than normal PoS interfaces because they are not meant to travel long distances. They can be manufactured with lasers that are weaker in strength and photodiodes that are not as sensitive as photodiodes required for long spans.

Figure 9-2 *ATM Aggregation over PoS Networks*



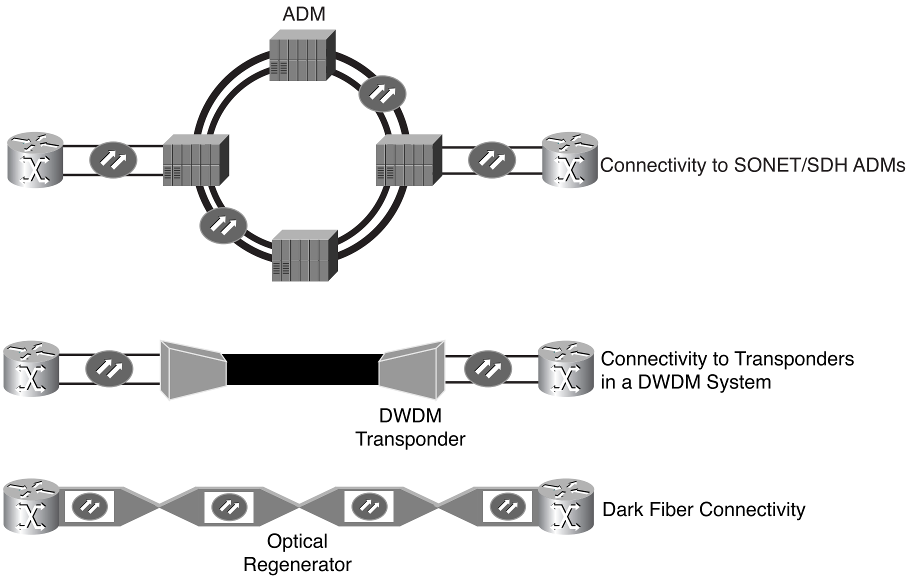
PoS Transport

PoS does not require SONET transport but works in tandem with such as a result of the SONET framing that PoS employs. Two PoS devices can be connected directly with duplex fiber. Because PoS interfaces are Layer 3 enabled, PoS interfaces are an example of an IP+Optical architecture.

Figure 9-3 displays three different ways in which PoS traffic can be transported. The three mechanisms are explained as follows:

- **Connectivity to SONET ADMs**—SONET circuits are provisioned as point-to-point circuits over SONET rings. Routers with PoS interfaces can be attached to SONET add/drop multiplexers (ADMs). As long as the proper number of STS are provisioned, the PoS interfaces will have connectivity. The PoS traffic is multiplexed with the other traffic that the SONET ADMs are carrying.
- **Connectivity to transponders in a DWDM system**—PoS traffic can be translated to a DWDM ITU-grid wavelength using a transponder. Most transponders support SONET framing. Through the DWDM system, 32 PoS circuits can be multiplexed onto one fiber.
- **Dark-fiber connectivity**—PoS interface can be connected directly over dark fiber using PoS interfaces. Dark fiber is fiber that is leased from a service provider; the customer provides the source (Laser or LED)and destination (photodiode receiver). This process is normally referred to as *lighting the fiber*. Long spans can be accommodated through standard SONET regenerators that provide regeneration, reshaping, and retiming (3Rs) of the signal. The Cisco 15104 is an OC-48 SONET regenerator that fits this application.

Figure 9-3 *PoS Transport Options*


ADM
Connectivity to SONET/SDH ADMs
Connectivity to Transponders in a DWDM System
DWDM Transponder
Dark Fiber Connectivity
Optical Regenerator

Multiaccess Protocol over SONET

Multiaccess Protocol over SONET (MAPoS) is a high-speed, link-layer protocol that provides multiple access capability over SONET/SDH.

MAPoS is defined in RFCs 2171 and 2176. The MAPoS frame format is based on HDLC-like framing for PPP. MAPoS is a frame switch that allows multiple nodes to be connected in a star topology to form a LAN using MAPoS.

NOTE You can find all RFCs online at http://www.isi.edu/in-notes/rfc*xxx*.txt, where *xxxx* is the number of the RFC. If you do not know the number of the RFC, you can find it by doing a topic search at http://www.rfc-editor.org/rfcsearch.html.

MAPoS can be used to allow SONET connectivity directly to the desktop. MAPoS is much more costly than Ethernet connectivity and has not been deployed, albeit a little in European markets. Most ATM-to-the-desktop environments have migrated their infrastructures to Ethernet technologies. MAPoS will probably never gain the market acceptance that Ethernet has. With SPs looking for more ways to leverage the low cost of Ethernet, it would be unlikely that enterprise environments roll out expensive SONET interfaces to their desktops in place of the ubiquitous, cost-effective Ethernet interfaces they currently use.

Packet over SONET Operation and Specifications

The current Internet Engineering Task Force (IETF) PoS specification is RFC 2615 (PPP over SONET), which obsoletes RFC 1619. The PoS RFCs define the requirements that are needed to transport data packets through PoS across a SONET network. These requirements are summarized as follows:

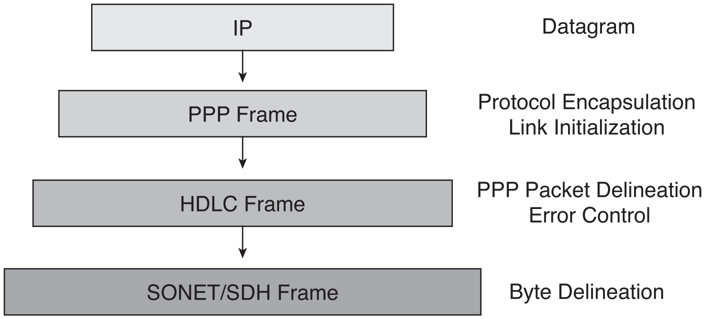
- **High-order containment**—PoS frames must be placed in the required synchronous transport signals used in SONET. An example of this is an OC-12 concatenated PoS interface. This interface requires an STS-12 circuit to contain the required payload of the PoS traffic.
- **Octet alignment**—This refers to the alignment of the data packet octet boundaries to the STS octet boundaries. An octet (byte) defines an arbitrary group of 8 bits. The word *byte* is defined as usually containing 8 bits. IBM used to define a byte as containing 7 bits. Although both byte and octet are used interchangeably, octet is a more accurate representation for 8 bits because its meaning is a series of eight.
- **Payload scrambling**—Scrambling is the process of encoding digital 1s and 0s onto a line in such a way that provides an adequate number for a 1s density requirement. The ANSI standard for T1 transmission requires an average density of 1s of 12.5 percent (a single 1 in 8 bits meets this requirement) with no more than 14 consecutive

0s for unframed signals and no more than 15 consecutive 0s for framed signals. The primary reason for enforcing a 1s density requirement is for timing recovery or network synchronization. However, other factors such as automatic-line-build-out (ALBO), equalization, and power usage are affected by 1s density. RFC 1619 inadvertently permitted malicious users to generate packets with bit patterns that could create SONET density synchronization problems and replication of the frame alignment. RFC 2615 provides a more secure mechanism for payload scrambling.

High-Order Containment

End stations at customer sites are predominantly TCP/IP-enabled devices. At the edge of the customer's network, the IP packet is encapsulated into a Layer 2 format that will be supported on the SP's network. The Layer 2 protocols supported by Cisco are PPP and Cisco HDLC, but the PoS standards specify PPP encapsulation for PoS interfaces. The Layer 2 PPP or Cisco HDLC frame information is encapsulated into a generic HDLC header (not Cisco proprietary HDLC) and placed into the appropriate SPE of the Whereas frame. This can be a confusing concept at first. Although HDLC and PPP are different, mutually exclusive Layer 2 protocols, HDLC is used as a SPE delimiter in the SONET frame. The encapsulation process of an IP packet to a SONET frame is illustrated in Figure 9-4.

Figure 9-4 *Encapsulating IP into a PoS Frame*



PPP Frame

RFC 1548 defines a PPP frame that contains the following three components:

- Protocol field
- Information field
- Padding field

The Protocol field is used because PPP was designed to be multiprotocol in nature. Multiprotocol encapsulations transport multiple protocols, including IP and IPX. The

Information field is the protocol data unit (PDU) transmitted, and can be from 0 to 64,000 bytes. The Padding field is used to pad the PPP frame if the Information field does not contain enough data. The Padding field might receive padding up to the maximum receive unit (MRU), which will fill the Information field. The default value for the MRU is 1500 octets but can be up to 64,000 octets if negotiated in the PPP implementation. It is the responsibility of the protocol to determine which bits are used as padding. You can find more information about the PPP protocol in RFC 1548 and RFC 1661 at www.ietf.org. Figure 9-5 illustrates the PPP in HDLC-like frame format.

Figure 9-5 *RFC 1662: PPP in HDLC-Like Framing*

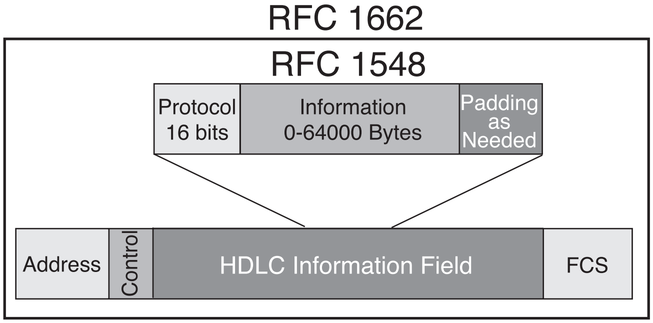
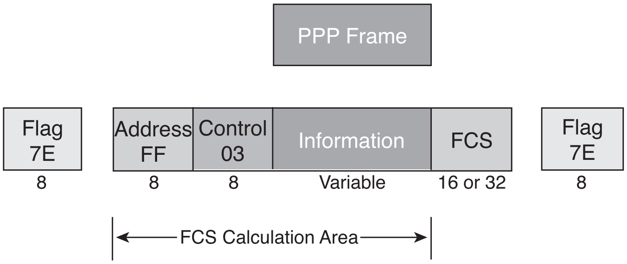


Figure 9-6 illustrates the values used in the PPP in the HDLC-like framing process. Notice that frame delimiters of hexadecimal 0x7E (126 in decimal) are used to denote the beginning and ending of a frame. The transmitting device generates flags as a time fill when there are no data packets.

Figure 9-6 *Packet over SONET Frame Information*



The Address field is always set to 0xFF (255) because every frame is a broadcast frame in PoS. There are only two ends of the point-to-point connection, and the frame always needs to get to the other side. There is no reason to have more than one address because there are no other addressable destinations. The Layer 2 mechanism is terminated at the other end of the link because PoS interfaces are Layer 3 enabled.

A Control field of 0x03 (3) is used to denote an HDLC frame. The Information field is where the PPP frame is inserted and is variable in nature due to MRU variability. A 16- or 32-bit frame check sequence (FCS) is used as a trailer to the frame. The FCS can be 16 or 32 bits long, but 32-bit CRCs are highly recommended due to the enhanced error recovery that is available using 32 bits. Most interfaces that run at speeds greater than OC-12 use FCS-32 as the default. The FCS is a configurable option, and FCS 32 is always recommended. The FCS field needs to match on both ends of the connection; otherwise, the Layer 2 protocol will never come up.

NOTE Although this book does not specifically deal with SDH, all Cisco PoS interface card framing can be changed from the default SONET framing to that of SDH.

Payload Scrambling

SONET has a default scrambler that was designed for voice transport. The 7-bit SONET scrambler is not well suited for data transport. Unlike voice signals, data transmissions might contain long streams of 0s or 1s. If the change from a 0 to 1 is not frequent enough, the receiving device can lose synchronization to the incoming signal. This can also cause signal-quality degradation resulting from distributed bit errors. The solution to this synchronization and bit error problem is to add an additional payload scrambler to the one normally found within SONET environments. This scrambler is applied only to the payload section, which contains the data. The SONET overhead bytes do not need this additional scrambling because they continue to use the existing 7-bit SONET scrambler. Certain overhead bytes, including the A1/A2 SONET framing bytes and the J0 section trace byte, are never scrambled with any type of scrambling.

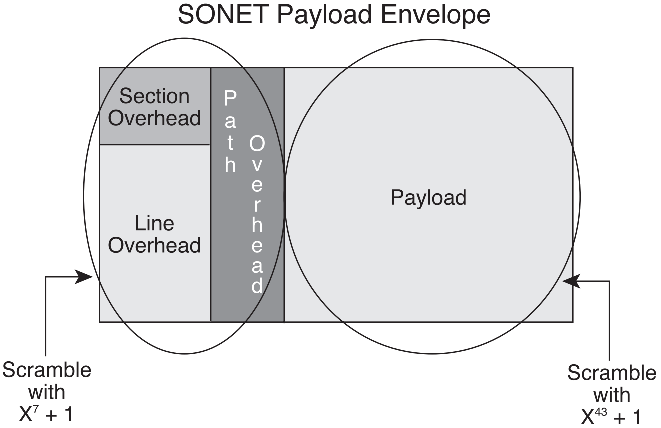
The two versions of scrambling that are supported by PoS are defined in the Telcordia GR-253 and ITU-T I.432 documents. The Telcordia GR-253 standard defines a basic $1 + x^{\wedge 6} + x^{\wedge 7}$ algorithm that scrambles the transport overhead of the SONET frame (with the exception of certain overhead bytes). This scrambler cannot be disabled and is adequate when the SONET frames carry phone calls in the payload.

The ITU-T I.432 standard defines an ATM-style scrambling. This scrambler uses a polynomial of $x^{\wedge 43 + 1}$ and is a self-synchronous scrambler, meaning that no state needs to be sent from the sender to the receiver. With this scrambler, only the data SPE of the SONET frame is scrambled.

The scrambling function is performed in hardware and does not affect performance in any way. Scrambling is performed directly in the framer application-specific integrated circuits (ASICs) on newer line cards and in a separate adjacent ASIC on older line cards. As technology evolves, more functionality is integrated into the same ASICs to lower the real estate (space) needs of hardware. Cisco supports port densities as large as 16 OC-3 ports on 1 Cisco 12000 series line card.

The path overhead C2 byte (path signal label) is used to instruct the receiving equipment that payload scrambling is turned on. If the traffic is carrying PPP with scrambling turned on, the value is set to a hexadecimal value of 0x16 (22). If scrambling is turned off, the original hexadecimal value of 0xCF (207) is used. Figure 9-7 shows the scrambling functions used in PoS environments.

Figure 9-7 *SONET RFC 2615 Payload Scrambler*



PoS Efficiencies

Overhead efficiency is a critical topic for SPs that charge for the amount of bandwidth capacity customers use. ATM was introduced to the SP market as the technology that would enable converged voice, video, and data traffic to reside on the same infrastructure (because of the intrinsic QoS parameters built in to the technology). The technology is widely used by Internet service providers (ISPs) today, but many of the original intentions behind ATM have not been used due to the complexity of configuring, selling, and maintaining such features. ISPs want to maximize their profits and minimize the costs associated with transporting IP over ATM.

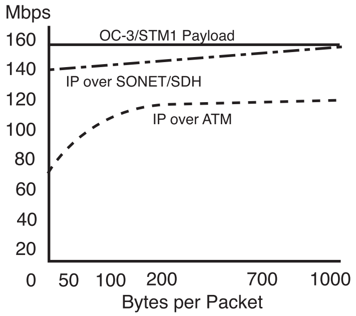
ATM uses fixed-sized ATM cells of 53 bytes. Each cell's composition includes 5 bytes of fixed overhead and 48 bytes of data. Depending on the ATM adaptation layer (AAL) used, the amount of AAL overhead can be as high as 4 bytes in addition to the 5 bytes of fixed overhead. This extra AAL overhead could result in as little as 44 usable bytes in a 53-byte cell. This equates to an approximate efficiency level of 83 percent (or 17-percent overhead). SONET's TOH and POH combined equal 36 bytes per the calculation that follows:

TOH (Transport Overhead)
9 rows × 3 columns = 27 bytes
POH (Path Overhead)
9 rows × 1 column = 9 bytes
TOH + POH = 36 bytes

The 36 bytes of overhead used in SONET represent approximately 4 percent of the total 810-byte STS frame. The ATM inefficiencies are further compounded when the variability of data sizes is calculated. Because of the nature of web pages, most Internet browsing traffic consists of many small-size packets. Most traffic that is generated originates from workstations connected to a LAN with Ethernet technology. The smallest PDU available in Ethernet networks is 46 bytes but can vary up to the maximum transmission unit (MTU) size of 1500 bytes. If a packet does not fall neatly on a cell boundary, the rest of the cell is padded.

In the case of a frame sent with a frame size of 64 bytes, two ATM cells are needed to transport the data. The first cell would be fully used at 48 bytes of payload (assuming that an AAL with no extra overhead is in use), and the second cell would be nearly empty with only 16 bytes of payload. This scenario results in a low efficiency level (approximately 43 percent). Figure 9-8 illustrates the PoS efficiency over ATM in both a line graph and table, which compares efficiency based on packet size.

Figure 9-8 *PoS Efficiencies Compared to ATM*



Packet Size (byte)	POS Eff. (SPE %)	ATM Eff. (SPE %)
64	86.8	43
128	94	69
256	97	75
512	98.6	85
1024	99.3	86
1518	99.5	88
2048	99.6	89
4352	99.8	89

PoS Network Designs

Resiliency is an important concern in SP networks. Outages result in lost revenue and might cause customers to cancel their service. SPs enter into Service Level Agreements (SLAs) with their customer. These SLAs guarantee certain levels of service. SLAs differ in many respects depending on the amount of risk the customer is willing to take. The more risk the customer is willing to take, the looser the SLA is and the cheaper the cost to the customer. The downside is that the customer is not guaranteed the same level of service as the customer who was not willing to take as much risk and paid more money for a stringent SLA.

PoS provides support of the optical 1+1 automatic protection switching (APS) mechanism. A customer desiring this level of protection orders two circuits from the SP: one for working traffic and one for protect traffic. SPs offer discounts for circuits that are used for protect traffic. The CPE router in this design could be a single point of failure, depending on how the circuit terminates at the CPE. Both circuits terminating on one line card of one router would result in a single point of failure from both the line card and router perspective.

A method of slightly higher resiliency is to still use one router, but use separate line cards for the working and protect circuits. This scenario provides fault tolerance in the case of a line card failure, but not a router failure. A higher level of fault tolerance might be achieved if each circuit terminates on a separate router.

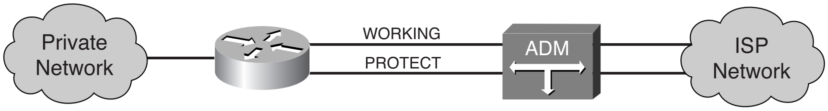
All of these survivability network designs are connected to one ADM at the service provider. APS 1+1 protection schemes are normally implemented per add/drop multiplexer. Ring failure is handled by the SP's robust SONET ring protection mechanisms. ADMs are carrier-class devices and must maintain a level of Five-Nines reliability. Five-Nines reliability refers to the amount of uptime a customer should expect from that network. Five-Nines reliability represents an uptime of 99.999 percent.

One Router

Figure 9-9 shows a design where there is one router at the customer premises with two optical interfaces used for APS 1+1 protection. Although a one-router CPE design does not provide the highest level of resiliency, this design does offer some advantages, including the following:

- No routing convergence upon failure of the working circuit or optical interfaces.
- 1+1 APS optical protection. Convergence time can be achieved in sub-60-ms time.
- Low-complexity network configuration.

Figure 9-9 *One-Router CPE Design*

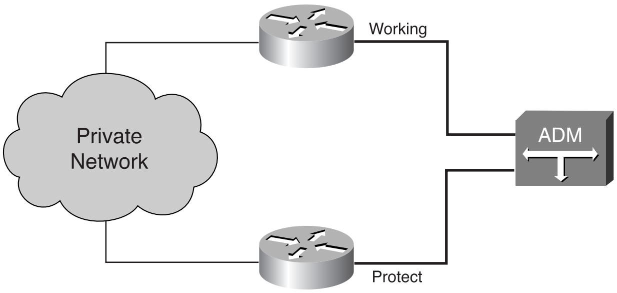


Two Routers

In a two-router design, each router has one optical connection to the SP's add/drop multiplexer. Fault tolerance has been increased with this design because the CPE router is no longer a single point of failure and the routers can be located in different areas of the building to facilitate fault tolerance associated with issues that could arise in isolated areas.

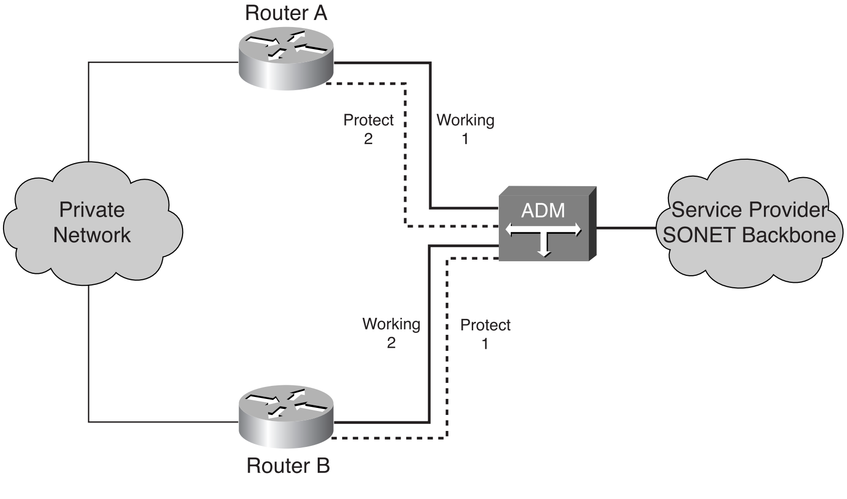
Figure 9-10 illustrates the two-router design philosophy. Although each router has one optical interface to the ADM, one link is working (active) while the other link is protecting (standby) the working link.

Figure 9-10 *Two-Router CPE Design*



Resources are wasted if only one router is actively forwarding traffic. To fully use both routers, you could use a design including four circuits. The design requires twice the number of interfaces and circuits, but this might still be cost advantageous depending on the amount of bandwidth required and the router hardware employed. Figure 9-11 displays an environment that includes two routers and four circuits in which both routers are in a working state for one circuit. The optical protection scheme is logically divided into APS protection groups that the routers monitor. A large router such as the Cisco 12000 can accommodate hundreds of APS groups.

Figure 9-11 *Two Routers with Four Circuits*



Advantages of having two PoS-connected routers include the following:

- Router redundancy in addition to circuit redundancy
- Load balancing of traffic

Disadvantages of having two routers connected to the single ADM in the SONET/SDH network include the following:

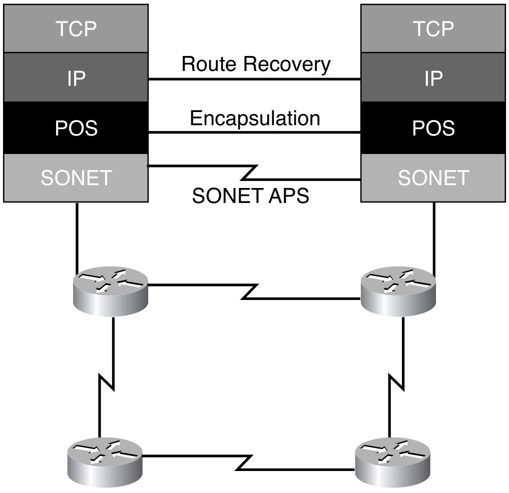
- **Convergence time**—The Layer 3 routing protocol must converge to optical circuit failure.
- **Complex network configuration**—APS groups.
- **Cost**—The costs associated with setting up and maintaining the design.

Failure recovery using two routers cannot achieve the sub-60-ms time that the one-router alternative offers. The Layer 3 routing protocol implemented in the infrastructure needs to reconverge around the failure. This is not an issue with one-router designs because both of the PoS interfaces on one router can have the same IP address with the PoS APS 1+1 configuration commands. This feature is allowable because only one of the interfaces is active at any one time.

PoS Protection Schemes

Packet over SONET protection uses the SONET APS 1+1 protection scheme. APS 1+1 looks at the K1 and K2 bytes of the SONET line overhead to determine whether issues exist with the SONET ring. A failure in the SONET network that affects the customer's working path causes a failover at the client site. This failover time occurs under 50 ms in the SONET Layer 1 network. The router interface uses a keepalive to determine whether the other side of the connection is alive. Keepalives are sent every 10 seconds by default, and the loss of 3 subsequent keepalives results in the interface going to an up/down state. After Layer 2 is lost, the Layer 3 routing protocol must converge around the link failure. Waiting for Layer 2 and Layer 3 to go through this procedure can take a long time (more than 30 seconds). Configuring the keepalives to 1 second lowers the convergence time to 3 seconds. Because PoS interfaces are Layer 3 implementations, the interfaces need to rely on a hierarchical error-recovery method such as that shown in Figure 9-12.

Figure 9-12 *PoS Hierarchical Error Recovery*



Layer 3 provides rerouting decisions during network failures to provide intelligent resiliency to the network. Layer 3 routing might be needed during a link failure if the Layer 3 IP address is changed. The Layer 3 IP address in the one-router design would be identical and there would never need to be Layer 3 routing protocol reconvergence. PoS interfaces in different routers require different IP addresses and always result in routing protocol reconvergence.

APS 1+1 Protection

SONET APS 1+1 is used for any PoS design that has more than one optical interface. APS provides optical protection during times of optical failure in the SP network. This information is carried over the K1 and K2 bytes of the SONET overhead. The CPE listens to the K1 and K2 bytes generated by the SONET network and generates K1 and K2 bytes when a failure occurs on the customer side. If the working interface of an APS 1+1 group fails, the protect interface can quickly assume its traffic load. The Layer 1 APS 1+1 recovery mechanism operates in 60 ms.

NOTE SONET rings have a 50-ms switchover time rather than the 60-ms switchover used in the Bellcore APS 1+1 specification. The APS 1+1 specification provides 10 ms for failure detection and 50 ms for switch initiation, which collectively equal 60 ms.

SONET APS works at Layer 1 providing switchover times significantly faster than any protocols operating at Layer 2 or 3.

The SONET protection mechanism used for PoS on Cisco products uses APS 1+1 with either unidirectional or bidirectional switching. (You can read more about 1+1 uni- and bidirectional switching in Chapter 3, “SONET Overview.”)

The SONET APS 1+1 architecture designates that there will be two circuits and each will carry the same traffic. One circuit is considered the working circuit; the other is the protect line. This differs from 1:1 or 1:*n* electrical-protection schemes because the backup equipment in electrical-protection schemes only carries traffic upon failure. The working and protect lines of APS 1+1 are both always transporting traffic. The receiving device(s) only process the traffic being received on the working circuit.

Protection mechanisms are more complex when circuits are terminated on different routers. The protection router must somehow be identified of the failure situation. An additional protocol is needed to provide for this signaling. This protocol is a Cisco proprietary mechanism called the Protection Group Protocol (PGP).

If a signal fail (SF) or a signal degrade (SD) condition is detected, the hardware switches from the working circuit to the protect circuit. APS 1+1 has reversionary capabilities allowing the hardware to switch back to the working circuit automatically when the original signal is restored for the configured time interval. The configurable reversion time is used to prevent the system from switching back to the working circuit if it is flapping (repeatedly going up and down). Flapping is sometimes referred to as *switch oscillation* and should be avoided at all costs so that the SP equipment can meet the SLAs. If the revertive option is not used, after a switch has moved to the protect circuit, the hardware does not automatically revert back to the working circuit. A system administrator must manually perform this function. Bidirectional switching is the default operation in Cisco routers. A circuit that automatically switches back to the original facility is called a *reversionary circuit*.

The K1/K2 bytes from the line overhead of the SONET frame indicate the current status of the APS connection and convey any requests for action. In standard APS, the two ends of the connection use this signaling channel to maintain synchronization.

With Cisco PoS, the working and protect channels are synchronized through an independent communications channel that is not part of the standard SONET APS system. This independent channel works whether the interfaces are on the same or different routers. This low-bandwidth connection is the Cisco PGP.

Cisco Protect Group Protocol (PGP)

PGP is the Cisco proprietary APS communication channel that is used between routers to complement APS 1+1 protection signaling. APS 1+1 is normally only done on the same router, but PGP enables this functionality to span multiple routers for added resiliency.

Performing APS 1+1 operation between routers creates some Layer 3 convergence issues. The standard Layer 2 mechanism used to determine whether an interface is down is the keepalive function. To accommodate fast reconvergence times, the keepalive update timer should be changed to 1 second and the hold timer changed to 3 seconds. PGP is the signaling channel used to inform the router with the protect facility about the failure. PGP operation closely resembles that of Cisco Hot Standby Router Protocol (HSRP) performing a heartbeat operation over a low-speed interface that tracks the status of certain ports. You can configure different protection groups to monitor multiple ports. The PGP protocol is a connectionless protocol that uses User Datagram Protocol (UDP) port 172 for message transfer. Figure 9-13 displays two routers that are configured in the same APS group. Notice that PGP updates are propagated bidirectionally between the working and protect routers to exchange information regarding the status of the PoS interface.

Figure 9-13 *Protection Group Protocol Operation*

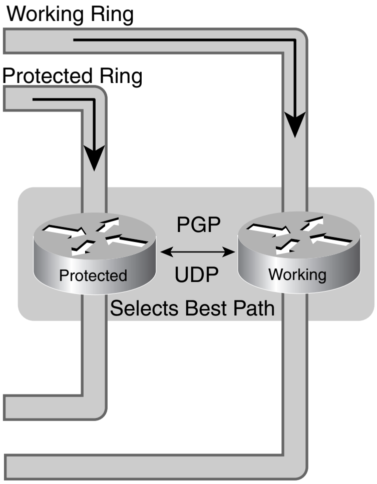
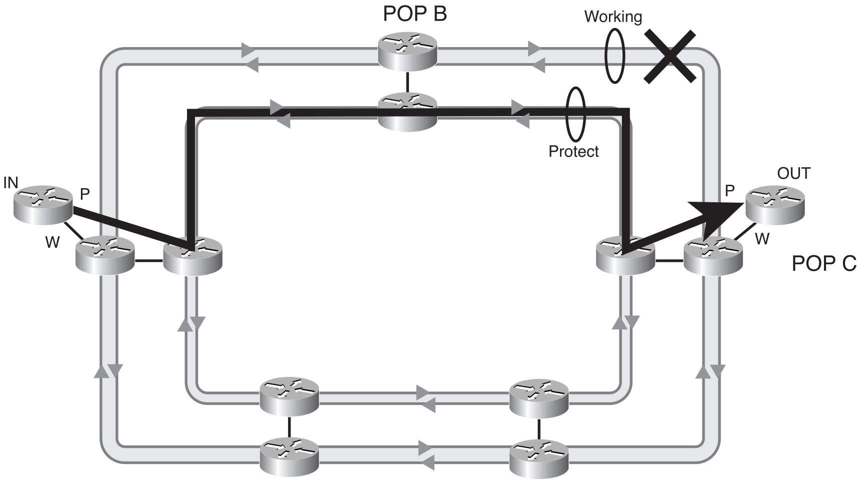


Figure 9-14 displays a network in which an outage occurs between POP B and POP C on the working facility. The routers at POP B and POP C will have knowledge of this outage through a loss of signal (LOS) condition, and PGP will notify the other router that it will now become the working interface. The other routers in the network will learn of this occurrence through the K1/K2 byte signaling occurring throughout the network.

Figure 9-14 *PGP Link Selection*



PoS Convergence

Convergence time is the amount of time required for all routers in a network to learn of changes in the network topology. Routers must propagate new route information from one end of the network to the other. Routing protocols are implemented to exchange this information. The routing protocol implemented should provide an ample amount of scalability to meet the future needs of the networks used in the environment. The faster the routing protocol can converge, the less downtime that will occur.

Scalable IP network routing protocols, such as Open Shortest Path First (OSPF), Integrated IS-IS, and Border Gateway Protocol (BGP), are responsible for recovering from error conditions in the network. Although the SONET APS 1+1 protection switching mechanisms guarantee a restoration time of 60 ms, the PoS interfaces are Layer 3 implementations and require some deal of routing protocol convergence. Typical convergence times for scalable routing protocols are several seconds or more depending on the environment and routing protocol design.

Figure 9-15 displays a design in which one router is used for the PoS interfaces. With this design, both of the PoS interfaces in the router can be configured with the same IP address. If a failure occurs, the router can perform switchover in the APS 1+1 switchover time of 60 ms. The Layer 3 routing protocol has not changed in any way on the LAN or WAN side of the router. The Layer 2 keepalive mechanism might not be aware of this switchover because it occurred in less than the lowest keepalive timer of 1 second. Regardless, three keepalives must be missed before an interface is determined as down.

Figure 9-15 *1-Router APS 1+1 Convergence*

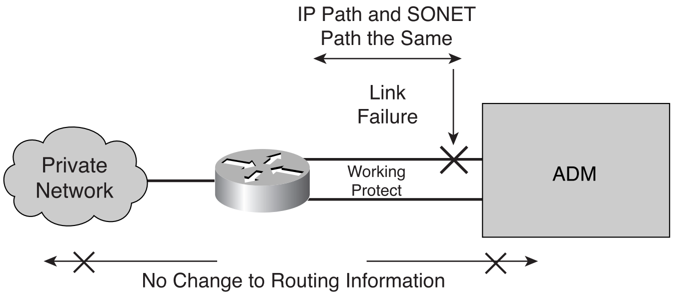
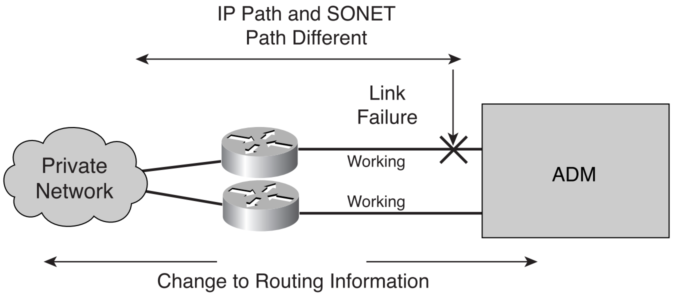


Figure 9-16 displays an environment that requires a higher degree of fault tolerance. This design uses two routers to implement the APS 1+1 group to protect the design from a router failure. The added resiliency creates some Layer 3 convergence issues because the interfaces used cannot have the same IP address if they reside on different physical routers. When the failure occurs, PGP is used to determine that the working interface has gone down, and the protect interface takes over. After this switchover has occurred, the Layer 3 routing protocol must communicate this information on both the LAN and WAN side so that the end to end network learns of the failure and solution. It is best to use HSRP on the LAN side if the PoS routers represent the default gateways out of the network. HSRP update and dead timers should be configured to match those of PGP.

Figure 9-16 *2-Router APS 1+1 Convergence*

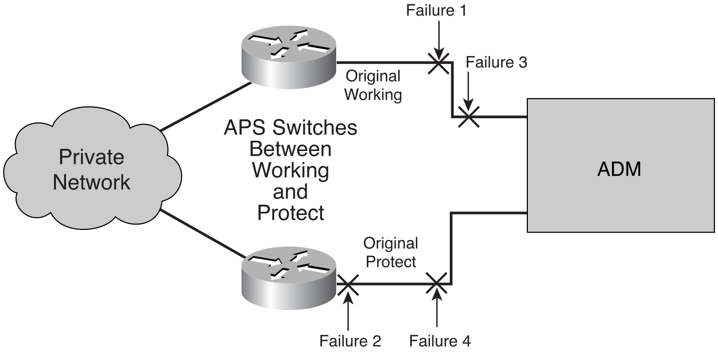

IP Path and SONET Path Different
Link Failure
Private Network
Working
Working
ADM
Change to Routing Information

Flapping

Flapping is the operation of a transmission line regularly transitioning from an up/up to an up/down state in a short period of time. Intermittent failures can result in the APS protection mechanism switching between the working and protection traffic repeatedly, causing many fluctuations in the network. If a two-router PoS model is implemented, the Layer 3 routing protocols will flap, too. You can see this issue in Figure 9-17.

APS switches traffic upon failures, but the routing protocol must send out routing updates. If another failure happens (Failure 2), the failure results in another APS switchover and more routing updates. Subsequent failures (Failures 3 and 4) repeat the process. The result of this flapping is that the network could end up spending all the time sending routing updates and reconverging around repeating failures instead of sending data across the network.

Figure 9-17 *Flapping in a 2-Router PoS Design*

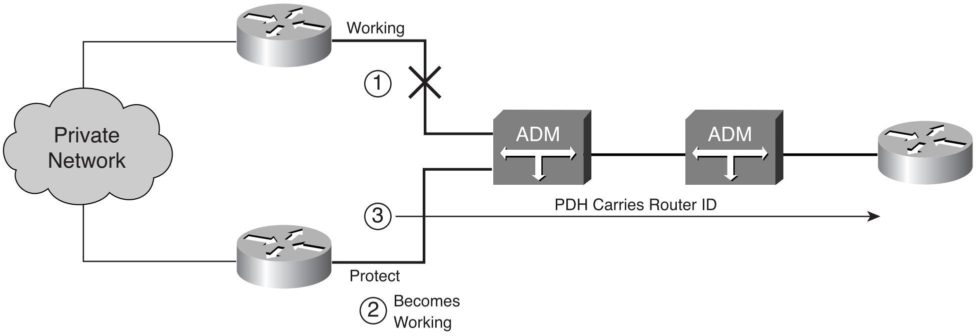


The issue is manageable by tweaking the reversion timer to a time greater than that necessary for the Layer 3 routing protocol to converge. The interfaces would not bring down the network because they must be stable in that amount of time before any switchover will take place.

PoS Reflector Mode

PoS Reflector mode is a process that is used to inform the remote router of a change in the network topology due to a line failure. Figure 9-18 displays an environment with two routers where a failure has occurred on the working line. As soon as the protect router receives information of the down interface through PGP, the protect router initiates a packet to the other side of the connection to speed up convergence. The packet contains the router ID information needed by the routing protocol to create the new Layer 3 adjacency. The remote router can now change the IP adjacency information immediately and reduce the convergence time dramatically.

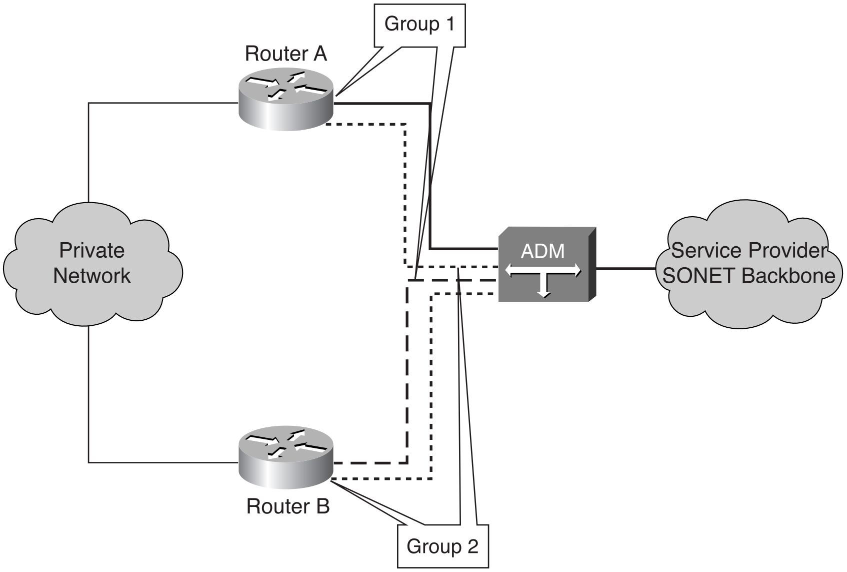
Figure 9-18 *PoS Reflector Mode*



Load Balancing

Load balancing refers to the capability to have traffic traverse two separate paths simultaneously to maximize the resources at the site. Load balancing is possible in a PoS APS 1+1 environment where four circuits are present. APS groups are configured on each router. One router is the working router for Group 1, and the other router is the working router for Group 2. Each of these routers protects each other using the PGP mechanism to alert the other side of failures. Figure 9-19 shows this design. You can use Multigroup HSRP (MHSRP) on the LAN side to actively forward traffic to both of these devices while providing the resiliency necessary. Layer 3 convergence is an end-to-end solution.

Figure 9-19 *PoS Load Balancing*



Alarms and PoS

Customers want to be notified of problems and errors that occur on their lines. PoS uses the same alarming of that used for SONET alarm reporting. The information that is carried in the overhead bytes of the Section, Line, and Path overhead layers are used by PoS to determine and report errors. This includes such items as the following:

- **Loss of signal (LOS)**—Signal failure due to a loss of light on the receive interface. A loss of light can also be thought of as receiving an all-0s pattern before descrambling. A downstream AIS should be sent when an LOS is detected.
- **Loss of frame (LOF)**—Issue created by receiving A1 and A2 bytes that do not indicate the 2-byte code of F628 in hexadecimal. An LOF condition is registered after no valid framing information has been received in 3 ms. The receipt of two subsequent valid A1/A2 frames clears this condition. A line alarm indication signal (AIS) must be sent downstream when this condition occurs.
- **Bit interleaved parity (BIP) errors**—BIP-3 errors occur at the path layer. The PoS interface is a path terminating equipment (PTE) device. The B3 byte carries the path parity errors in this byte.
- **Loss of pointer (LOP)**—When a pointer processor cannot obtain a valid pointer condition, an LOP state is declared, and a downstream AIS must be sent. Recall that the H1 through H3 bytes of the LOH are used for the pointer functionality.

Threshold registers record all the normal SONET counters for errors that occurred over the past 15 minutes and past 24 hours. You can view these by using IOS **show** commands. When the threshold register exceeds the threshold register settings, a threshold crossing alarm (TCA) indication occurs, meaning the device needs to notify the management station of the alarm.

Summary

This chapter covered PoS operation, encapsulation, protection, and convergence. You should be able to describe the most popular uses for PoS. You should be familiar with the PoS frame structure and encapsulation and be able to describe the efficiencies of PoS over other technologies. PoS design models provide advantages in the area of resiliency but need the PGP mechanism to decrease convergence times. You can achieve load balancing in PoS networks by creating multiple APS groups and using MHSRP on the LAN side of the connections. PoS is implemented as a point-to-point technology.

Review Questions

1 True or False: Packet over SONET was developed because there was no other way to transport data over a SONET network.

A True

B False

2 PoS can be directly encapsulated onto the network media. Which of the following is not a method for connecting PoS to network media?

A Connectivity to SONET/SDH ADMs

B Connectivity to transponders in a DWDM system

C Dark-fiber connectivity

D ATM connectivity

3 Which of the following are the three requirements for data to be successfully transported over SONET/SDH?

A The use of high-order containment is required.

B The PoS frames must be placed inside of the SONET containers aligned on frame boundaries.

C The $x^{\wedge 43 + 1}$ scrambler must be used in addition to SONET native scrambling.

D The PoS frames must be placed inside of the SONET containers aligned on the octet boundaries.

4 What is the hex value of an HDLC delimiter flag byte?

A 0x7D

B 0xE7

C 0x7E

D 0xFF

E 0x03

5 What does the protocol field inside of the PPP frame indicates?

A The protocol that is carrying the PPP frame

B The protocol used to decode the FCS field

C The protocol used to detect the number of padding bytes found in this frame

D The protocol used to format the data in the Information field

E None of the above

6 The C2 byte value of a PoS interface that is using the payload-scrambling function is set to which two of the following values?

A 0xCF%

B 0x16

C 22

D 207

E 0xFF

7 It is recommended that the Layer 3 protocol and the SONET protocol configurations should ______.

A Have both Layer 3 and SONET in a bidirectional ring configuration

B Have Layer 3 in a point-to-point configuration and SONET in a bidirectional ring configuration

C Both be in a point-to-multipoint configuration

D Both be in a point-to-point configuration

8 HDLC frames in a PoS environment contain which four fields?

A Address, Data, Destination, and Frame Check Sequence

B Location, Data, Destination, and Protocol

C Location, Control, Information, and Protocol

D Address, Control, Information, and Frame Check Sequence

9 What is the purpose of PGP?

A Transport data packets across SONET/SDH links

B Overcome routing problems between Layer 3 and the SONET network layer

C Reliable end-to-end communication and error-recovery procedures

D Achieve adequate transparency, protection against malicious attacks, and enough zero-to-one transitions to maintain synchronization between adjacent SONET/SDH devices

10 What is the ideal configuration for APS 1+1 to reduce the need for routing updates due to a failure?

A One SONET line between the private network router and the service provider ADM

B Two SONET lines between one private network router and the service provider ADM

C Two routers with one line each to the service provider ADM

D PoS Reflector mode

11 PoS Reflector mode is used for what purpose?

A By the working router to keep the distant router up to date

B By the protect router to notify the distant router when it takes over for the working router

C By the Layer 3 protocol to send routing updates

D To send AIS downstream

12 In which three ways can you interconnect PoS interfaces?

A SONET

B Dark fiber

C Gigabit Ethernet

D DWDM

E Bidirectional path switched rings (BPSRs)

13 One of the advantages to PoS is that when there is a network failure, ___________ can restore the network connection before the Layer 3 routing protocol even realizes that there is a problem.

A ATM

B SDH

C APS

D IPS