A Tutorial on SONET/SDH Automatic Protection Switching (APS)

Technology White Paper

Steve Gorshe
Principal Engineer

Issue 1.0: February, 2005

© 2005 PMC-Sierra, Inc.
Abstract

The simplest class of mechanism for network survivability in the event of failure on a network element or link is automatic protection switching (APS). APS schemes involve reserving a protection channel (dedicated or shared) with the same capacity as the channel or facility to be protected. This white paper provides an overview of the different types of protection switching available with SONET/SDH networks, focusing on technologies that automatically work at the level of either the SONET (SDH) Line (Multiplex Section) or Path.

While SONET/SDH protection began with the relatively simple point-to-point line protection scheme, it has proved to be very extensible in terms of topologies (e.g., rings) and flexibility (e.g., Link Capacity Adjustment Scheme (LCAS) service restoration). The 1+1, 1:1, Unidirectional Path Switched Ring (UPSR), and Bi-Directional Line Switched Ring (BLSR) technologies each have an important role in the network. LCAS-based restoration shows promise for increasing the cost effectiveness of many packet-based services. As readers of this white paper become familiar with the processing requirements for APS and LCAS service restoration, they will appreciate the enhanced functionality in the PMC-Sierra CHESS III, Arrow, and ADM devices which greatly ease APS implementations.

About the Author

Steve Gorshe, Ph.D. is a Principal Engineer in the Product Research Group and oversees ICs for SONET, optical transmission and access systems.

Currently Steve is a senior member of the IEEE and co-editor for the IEEE Communications magazine’s Broadband Access Series. He is the chief editor for the ANSI T1X1 Subcommittee, which is responsible for SONET and optical network interface standards. He is a recent recipient of the Committee T1 Alvin Lai Outstanding Achievement Award for his standards work and has been a technical editor for T1.105, T1.105.01, T1.105.02, and T1.105.07 within the SONET standard series as well as the ITU-T G.7041 (GFP) G.7043 (Virtual concatenation of PDH signals), G.8040 (GFP mapping into PDH signals), and G.8011.1 (Ethernet Private Line Service) recommendations. He has 26 patents issued or pending and several published papers.

Revision History

<table>
<thead>
<tr>
<th>Issue No.</th>
<th>Issue Date</th>
<th>Details of Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>February, 2005</td>
<td>Document created</td>
</tr>
</tbody>
</table>
Contents

Abstract .............................................................................................................................. 1
About the Author ............................................................................................................. 1
Revision History ............................................................................................................ 1
Contents ............................................................................................................................. 2
List of Figures ................................................................................................................... 3
List of Tables ..................................................................................................................... 3
Preface ............................................................................................................................... 4

1 Introduction ............................................................................................................... 5

2 1+1 and 1:N SONET APS .......................................................................................... 6
   2.1 1+1 Protection .................................................................................................... 8
   2.2 1:N Protection ..................................................................................................... 8

3 Subnetwork Connection Protection (SNCP) ........................................................ 10

4 Survivable Rings ....................................................................................................... 11
   4.1 Ring Classifications .......................................................................................... 11
   4.2 SONET Unidirectional Path Switched Rings (UPSR) ........................................ 13
   4.3 SONET Bi-directional Line Switched Rings (BLSR) and SDH
       Multiplexed Section Shared Protection Ring (MS-SPRing).............................. 14
       4.3.1 BLSR / MS-SPRing Protection Switching ............................................ 15
       4.3.2 Prevention Against Traffic Misconnection ........................................... 18
       4.3.3 Extra Traffic and Non-Preamptible Unprotected Traffic ......................... 19
   4.4 Ring Protection Processing Challenge ............................................................. 19
       4.4.1 TFI-5 Interface ..................................................................................... 20
   4.5 Interconnected Rings ........................................................................................ 20
       4.5.1 Interconnection with UPSRs ................................................................. 21
       4.5.2 Interconnection with BLSRs................................................................. 21
       4.5.3 Interconnection with Intermediate Multiplexing and Payload
           Defect Indication (PDI)............................................................................. 22
       4.5.4 Other Ring Interconnection Considerations ........................................... 23

5 Partial Protection Using the Link Capacity Adjustment Scheme (LCAS) .......... 24

6 Conclusions ............................................................................................................. 26

7 References ............................................................................................................... 27

8 Notes ........................................................................................................................ 28
List of Figures

Figure 1  Linear APS: (a) Basic protocol steps, (b) 1+1, (c) 1:n example ....................... 7
Figure 2  Ring routing examples .................................................................................... 12
Figure 3  UPSR illustration............................................................................................. 13
Figure 4  Two-fiber BLSR illustration with cable failure ................................................. 15
Figure 5  Two-fiber BLSR illustration with node failure .................................................. 17
Figure 6  Potential traffic misconnection problem that a BLSR must handle................. 18
Figure 7  Dual Node Interconnection of UPSRs ............................................................ 21
Figure 8  Drop-and-continue and selection functions in interconnection of
BLSRs (only one traffic direction shown)........................................................................ 22
Figure 9  Ring interconnection examples......................................................................... 22
Figure 10 LCAS fallback to reduced bandwidth due to failure affecting two out of
three members of the VCG............................................................................................. 25

List of Tables

Table 1  SONET/SDH Byte K1 and K2 definitions for linear APS (from ANSI
T1.105.01).................................................................................................................... 9
Table 2  Byte K1 and K2 definitions for a SONET BLSR (and SDH MS-SPRing) .......... 16
Preface

The goal of this white paper is to provide the reader with an overview of the different types of protection switching available with SONET/SDH networks. As such, it is a companion to the white paper tutorial on SONET/SDH [8]. The focus here is on those technologies that automatically work at the level of either the SONET (SDH) Line (Multiplex Section) or Path. Mesh-based restoration schemes are not covered here since they are typically not accomplished through automatic responses to simple signaling. The reader who is interested in mesh-based approaches is referred to [7]. The white paper provides detail at the level of the bit and byte definitions for the SONET/SDH APS channels. The description of the SONET Bidirectional Line Switch Ring (BLSR) / SDH Multiplex Section Shared Protection Ring (MS-SPRING) provides a thorough overview of the protocol, but omits the specifics of the signaling state machine that can be found in the applicable standards [2] [4].

The information in this white paper has also been adapted to form part of the following textbook: M. Elanti, S. Gorshe, L. Raman, and W. Grover, Next Generation Transport Networks – Data, Management, and Control Plane Technologies, Springer, 2005.

The author wishes to thank Wayne D. Grover of TR Labs / University of Alberta for his comments and additions to the material in this white paper.
1 Introduction

The simplest class of mechanism for network survivability in the event of failure on a network element or link is automatic protection switching (APS). APS schemes involve reserving a protection channel (dedicated or shared) with the same capacity as the channel or facility to be protected. Different APS techniques are characterized by the following criteria:

- The topology, either linear or ring;
- Whether the protection channel carries a backup copy of the traffic permanently or only when requested for protection;
- Whether the protection channel is shared among working channels that may potentially need protection;
- Whether both directions of transmission switch (bi-directional switching) to protection channels when a failure occurs in one direction or only the affected direction switches (unidirectional switching); and
- Whether the network automatically reverts traffic back to the working channels after they have been repaired/restored (revertive switching) or continues to use the protection channel after the repair/restoration (non-revertive switching).

In so-called “linear protection” the entity to be protected follows a point-to-point route for that layer of the network where the end-nodes of the protected path segment are different. “Linear” distinguishes this type of APS scheme from ring schemes where the protected path closes on itself at the end-nodes. SONET/SDH APS is implemented at the Line/Multiplex Section layer. Ring protection uses an alternative path around the ring in the opposite direction to the one affected by the failure.

The standards for ring and APS techniques are still evolving. ANSI T1.105.01 [2] covers SONET linear and ring APS and ITU-T Recommendation G.841 covers linear and ring protection for SDH networks [3] [4]. The ITU-T has recently begun a project to define a general set of recommendations for protection that are independent of the underlying transport technology. The first of these is G.808.1 [5] for linear networks. G.808.2 will cover ring networks. The protection recommendations that will be developed for G.709 OTN networks (and later revisions of G.841) will reference these generic protection recommendations for common concepts rather than repeating them in each technology-specific recommendation.
2 1+1 and 1:N SONET APS

In any APS scheme, the network element (NE) that detects the fault condition also initiates the protection switching action and is referred to as the tail-end node. The node at the other end of the protected link is referred to as the head-end node. Note that these definitions apply separately to each direction of transmission. The head-end node’s main task is to electrically split the affected working signal (i.e., make an electrical duplicate copy of the signal) and feed this bridged signal into the standby protection channel (while continuing to feed the working channel as well). This is referred to as the head-end bridge function. In 1+1 APS, the bridge is always present. In 1:1 or 1:N, the tail-end signals upstream to request the bridge upon failure. In SONET/SDH, this signaling information is communicated in the K1 and K2 overhead bytes; these two bytes constitute the SONET/SDH APS signaling channel. G.709 OTN has also reserved overhead bytes for implementing an APS signaling channel. The typical criteria for initiating APS are:

- Detection of a failure (e.g., loss of signal (LOS), loss of framing (LOF)),
- Signal Fail (excessive bit error rate (BER)),
- Degraded Signal (relatively high BER), or
- Externally initiated commands from the craft or OSS (e.g., manual switches or forced switches).

These criteria form a hierarchy of priorities when multiple channels or conditions compete for access to the protection resources. Forced switches are the highest priority, and would typically be performed when the craft needs to lockout a facility during maintenance or upgrades when other fault or failure conditions could be affecting other working channels. Failures are the next highest priority. In practice, high BER is treated as a failure. Manual switches are those initiated for maintenance purposes when the craftsperson wants the network to still automatically protect itself in the event of a real problem. Manual switches have the lowest priority.

In addition to being used by the tail-end to request signal bridging actions from the head-end NE, the APS signaling channel is also used to communicate status information between the head- and tail-ends. In SONET/SDH, the K1 byte primarily carries the tail-end requests and/or status, and the K2 byte communicates the head-end status. It should be noted that the APS communication is conducted over the protection channel rather than over the working channel. Because the system cannot rely on the working channel for communication when it has failed, the receiver ignores the APS signaling on the working channel. To test the protection control functionality, SONET/SDH standards also include a protection switch Exerciser mode. The Exerciser function performs all stages of protection switching except for the final action of bridging or switching the actual traffic, since this would put a ‘hit’ on the operating signal. The Exerciser function is performed routinely as a background maintenance activity to avoid the implications of a “silent failure” in the redundant standby equipment and protection switching logic.

---

1 A protection lockout command also allows the craftsperson to prevent traffic from being switched onto the protection channels. Protection lockout typically has the highest priority.
Figure 1  Linear APS: (a) Basic protocol steps, (b) 1+1, (c) 1:n example
2.1 1+1 Protection

1+1 protection switching is by far the simplest type of APS. In 1+1 there is a permanent head-end bridge of the working channel onto the protection channel. Typically, in SONET/SDH the channel here is an entire OC-n rate line signal. As illustrated in Figure 1(b), when the receiving NE detects that the traffic arriving from the protection channel is healthier than the traffic from the working channel, it switches to taking its traffic from the protection channel. 1+1 protection can be either unidirectional or bi-directional. In unidirectional 1+1 protection switching there is no need for the tail-end to communicate with the head-end (i.e., the tail-end simply chooses the best signal without informing the head-end). On the other hand, in bi-directional 1+1 protection switching, the tail-end needs to inform the head-end so that the head-end can initiate protection switching steps for the other direction of transmission that result in the head-end also taking its traffic from the protection channel. In non-revertive operation, the distinction between the working and protection channels disappears, except for the default starting state of the network.

In 1:N APS, a single protection or standby line-rate transmission system is shared by N working systems as illustrated in Figure 1(c).

2.2 1:N Protection

To support the shared access to a standby system, no head-end bridge is established until failure occurs. Signaling is then needed to specify the requested channel and request the head-end bridge. To this end, SONET/SDH standards define the K1 and K2 byte APS channel definitions in Table 1. 1:N APS greatly enhances system availability against failures of single fibers, transmitters or receivers—anything that affects only one working channel at a time. Following a failure of a working system, the tail-end informs the head-end that it needs its working-line signal to be bridged into the protection line. To do this the tail end originates K1 with a reference to the working line number and the fault type (SD, SF etc). The head-end NE then checks the current status of the protection system and the priority of the conditions (K1 bits 1-4) for each working system requesting protection.

As the K1 byte encoding in Table 1 illustrates, the maximum number of working channels is 14 plus one protection channel. The K1 byte encodes channel number zero to indicate protection line and fifteen to indicate extra traffic. Because 1:N protection has a protection line that is normally idle, it can be used to carry Extra Traffic.

Obviously 1:N protection is much more capacity-efficient than 1+1. The limitation is, however, that when applied to several fibers or wavelengths to form a 1:N APS system on the same cable, there is no survivability against cable cuts. There is thus an excellent internal system availability enhancement but no means of network restoration unless all working channels and the protection system follow disjoint physical routes. 1:N protection has therefore, not seen much use in practice. Its real significance, however, has turned out to be the simple extension of 1:N SONET linear APS into the SONET BLSR ring configuration where the signaling for channel number is adapted to apply to node number instead, in a closed ring configuration which then withstands cable cuts as well as single-channel failures.
### Table 1  SONET/SDH Byte K1 and K2 definitions for linear APS (from ANSI T1.105.01)

<table>
<thead>
<tr>
<th>Bits 1-4</th>
<th>Byte K1</th>
<th>Byte K2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111</td>
<td>Lockout of protection</td>
<td>Bits 1-4</td>
</tr>
<tr>
<td>1110</td>
<td>Forced Switch</td>
<td></td>
</tr>
<tr>
<td>1101</td>
<td>Signal fail – high priority (not used in 1 + 1)</td>
<td></td>
</tr>
<tr>
<td>1100</td>
<td>Signal fail – low priority</td>
<td></td>
</tr>
<tr>
<td>1011</td>
<td>Signal degrade – high priority (not used in 1 + 1)</td>
<td>Bits 5-8</td>
</tr>
<tr>
<td>1010</td>
<td>Signal degrade – low priority</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>Manual switch</td>
<td></td>
</tr>
<tr>
<td>0111</td>
<td>(not used)</td>
<td></td>
</tr>
<tr>
<td>0100</td>
<td>Exerciser</td>
<td></td>
</tr>
<tr>
<td>0011</td>
<td>(not used)</td>
<td></td>
</tr>
<tr>
<td>0010</td>
<td>Provisioned for 1:n mode</td>
<td></td>
</tr>
<tr>
<td>0000</td>
<td>No request</td>
<td></td>
</tr>
</tbody>
</table>

**NOTES –**
- The Lockout of Protection switch priority uses bits 5–8 = 0000.
- For Signal Fail and Signal Degrade only, bit 4 indicates the priority assigned to the working channel requesting switch action.
- The Exerciser function may not exist for certain protection switching systems.
- Reverse Request assumes the priority of the request to which it is responding.


3 Subnetwork Connection Protection (SNCP)

Subnetwork connection protection (SNCP) can be thought of as 1+1 APS applied end-to-end over an entire network, and implemented at the tributary signal level, as opposed to the whole OC-n line-rate signal level. For example, it could be used over the entire customer path (trail) from ingress to egress from the SONET/SDH network. This type of user-level 1+1 APS is a commonly employed technique in Europe, where the line-rate ring protection schemes that SONET/SDH can also support (i.e., BLSR/MS-SPRing) and UPSR are rarely used. The concept is identical to 1+1 APS but it would be implemented at, say, an STM-1 (STS-3c) tributary signal level, end-to-end over an entire network. This is in contrast to 1+1 APS applied at the entire OC-192 line rate of a specific transmission system facility, and routing customer tributary signals over the protected facility. At the egress of the subnetwork on which an SNCP arrangement is established, a selector evaluates the two signals that it receives over the two paths and chooses the best signal.

To avoid unnecessary or spurious protection switching in the presence of bit errors on both paths, a switch will typically only occur when the quality of the alternate path exceeds that of the current working path by some threshold (e.g., an order of magnitude better BER). SNCP allows protection of multiple subnetworks along an end-to-end path. The UPSR ring protection mechanism (to follow) can be considered as a special case of SNCP. North American SONET networks typically use UPSR for access and BLSR for transport rather than SNCP.2

2 SDH has an overhead channel reserved for “trail” protection. SNCP can take place at any of the subnetwork boundaries, while trail protection takes place only at the trail termination points. Otherwise the two are conceptually the same.
4 Survivable Rings

Ring-based protection schemes have enjoyed enormous popularity, especially in North America, for several reasons: Technically rings are enhancements of APS technology, and were thus relatively easily developed and standardized based on extensions to the SONET APS signaling protocol. Rings are thus the closest technology to prior generations of transmission systems that used 1+1 and 1:N APS and were thus relatively easy to develop and quick to gain acceptance by the transmission engineering community in the telcos. Secondly, rings address both the need for single-channel protection switching and by virtue of the ring topology, also protect against cable cuts. Rings also provide a system design that collects demands together to exploit the economy-of-scale in transmission technology (for instance an OC-192 ring may cost only two to three times what an OC-48 system at ¼ the capacity costs). In addition rings became available just when the need in industry reached crisis proportions after several spectacular cable cuts in the 1990s. At the time, rings were also perceived as being much simpler to understand and operate than distributed mesh-restorable schemes. Although more efficient mesh-based survivability architectures were then under study for SONET, rings filled the void as the first reasonable, standardized, and available solution for transport survivability and hence were rapidly and extensively deployed. A conservative estimate is that from 1990 to 2004, more than 100,000 SONET ring ADM terminals (both BLSR and UPSR and over the full range of OC-n rates) were deployed in North America alone. Rings are therefore fundamentally important in transport survivability and will be present as a legacy technology for a long time. One drawback with rings is that despite the impression of simplicity when considering a single ring, practical networks employing multiple interconnected rings have turned out to be extremely complex to design, operate, and grow, and to be quite inefficient in overall capacity usage. They are also relatively inflexible to changes in demand pattern. These are some of the reasons for growing interest in the mesh-based alternatives that follow.

4.1 Ring Classifications

There are two main categories of survivable ring. Path switched rings perform their protection at the path level, and line switched rings perform their protection at the line level (equivalently the “multiplexed section” level in ITU/SDH terminology). In principle a ring can support either unidirectional or bidirectional routing as illustrated in Figure 2, and either type of protection action (path or line). In practice, however, only two of the four possibilities have been developed as SONET rings. The SONET Unidirectional Path Switched Ring (UPSR) uses unidirectional routing and path level protection. The SONET Bidirectional Line Switched Rings (BLSR) and Multiplexed Section Shared Protection Ring (MS-SPRing) in SDH use bi-directional routing and line-level loopback protection. UPSRs are inherently “two-fiber” structures but BLSRs are defined in both “two-fiber” and “four-fiber” variants, to be explained shortly.

---

3 A reviewer of this material advised that in practice carriers aim for four times capacity at 2.5 times the cost, for a 40% reduction in cost per unit capacity for the larger OC-n system.
Consider the example of traffic exchanged between two ring nodes A and B in Figure 2 to see how ring routing affects the protection switching. In unidirectional rings, the working path from A to B and the path from B to A are routed in the same direction around the ring (clockwise in this example). The A-to-B and B-to-A connections typically occupy the same time slot on all spans of the ring. For bidirectional rings, the same channels can potentially be used for different inter-node communications in different parts of the ring. In the bidirectional ring example, nodes A and B communicate with each other directly over counter-propagating channels via the shortest route between the nodes on the ring. This allows nodes C and E to communicate over the same time-slots on another portion of the ring. Such reuse of wavelengths or timeslots on different portions of the ring is known as spatial reuse and can provide a great increase in ring capacity.

A disadvantage of unidirectional routing is that it does not allow spatial reuse of time slots on the ring. Spatial reuse is only useful, however, if the demand is distributed in a general way between pairs of nodes on the ring. Commonly in access networks all of the ring traffic originates from, and is destined to a single node, which is the hub node accessing the wider outside world and core network. In this case there is no efficiency benefit to bidirectional routing because in this case the flows all add up such that the cross-sections of the spans next to the hub node require the same line capacity in a BLSR as in a UPSR for the same application. This is why access ring networks, which are typically arranged as a collection of remote nodes that are all logically connected to the same central office node, are virtually always based on UPSR rings.

For long haul networks, the individual facilities are typically long and expensive, which makes spatial reuse economically attractive. Hence, most backbone long haul networks use BLSR. For metro transport networks, however, the facility costs are relatively lower and BLSR equipment has typically been somewhat more expensive than UPSR equipment. As a result, BLSRs are deployed most often in long-haul networks and core metro networks, while UPSRs are the most common in access/aggregation networks.
4.2 **SONET Unidirectional Path Switched Rings (UPSR)**

A UPSR is illustrated in Figure 3. Each UPSR node bridges traffic that enters (is “added”) at its site onto channels in both ring directions. The receiving node then compares the quality of signal received from the two paths around the ring and chooses the best path based on measures such as BER, framing loss, or signal level. UPSRs are typically non-revertive, so other than an initial provisioning default, the two paths are not typically designated as working and protection paths. Notice that this is effectively a collection of tributary-level 1+1 APS arrangements, or equivalently SNCP setups, within the ring subnetwork. As with SNCP, any path being dropped from the ring (i.e., at its egress point) is monitored. If a BER threshold, synch loss, or signal loss alarms arise, a tail-end transfer switch to the alternate path will be performed.

**Figure 3   UPSR illustration**

The main virtue of the UPSR is its simplicity. Protection switching decisions are made locally on a per-path basis, by each receiving node independently, so there is no need for APS signaling between the switching node and the corresponding head-end node. UPSR technology is simple and low-cost, especially for the access network application. Imagine the cost of separate (even non-redundant) point-to-point access links at, say, the OC-3 rate, compared to the same set of accessing nodes sharing a single ring-structured OC-48. This is the economy of scale effect mentioned above. First cost, space, power, inventory and maintenance implications correlate strongly with the simple number of transmitter and receiver circuit packs required, and only secondarily to the actual speed of those circuit packs. So a UPSR using a single pair of OC-48 optics cards at each site can be considerably more economical than the equivalent capacity obtained with lower-speed cards. Outside the access network, however, the UPSR is notoriously inefficient in its use of capacity. Because each signal carried in a UPSR appears everywhere in each span of the ring, it follows fairly directly that the UPSR line rate must equal or exceed the sum of all the individual demands it serves. This is the penalty associated with not supporting spatial reuse of capacity as in BLSRs.
Another practical design challenge with UPSRs is that the receiving NE requires circuitry to process BER and other monitoring functions continually on both paths for every individual tributary signal that egresses the ring at its site. This information is required to choose the best quality signal to drop. This can require a surprisingly large amount of CPU capability and related power consumption and heat dissipation, especially if every UPSR NE is designed to possibly terminate (drop) all tributaries from the ring. The “hub” node on a fully loaded OC-12 VT accessed UPSR would have to monitor the performance of 672 VT1.5 paths, including calculating and comparing the BER under degraded signal conditions. In contrast, in the BLSR, only aggregate states of the entire line-signal need be monitored to activate protection switching.

4.3 SONET Bi-directional Line Switched Rings (BLSR) and SDH Multiplexed Section Shared Protection Ring (MS-SPRing)

Line switched rings in general seek to provide greater capacity efficiency and to avoid the complexity of looking at each individual path for both the working and protection channels on higher rate systems (i.e. \(\geq 622\) Mb/s). By switching at the line level, all paths carried over that line are protected simultaneously. This in turn allows the receiving node to only provide monitoring and termination circuits for the paths that it receives over the working line. While unidirectional line switched rings have been proposed, carriers have typically preferred to include bi-directional routing capability into line switched rings. Figure 4 illustrates a four node BLSR.

In a typical configuration, half the bandwidth in each direction around the ring is reserved for working traffic and half for protection traffic with a fixed one-to-one correspondence between working and a protection channels. (An exception to this is noted below for non-preemptible unprotected traffic.) In a two-fiber BLSR, the working traffic uses half the STS paths for working traffic and half for protection. In a four-fiber BLSR there is a dedicated working fiber and a corresponding protection fiber in each ring direction.

There are advantages of both four-fiber (4F) and two-fiber (2F) BLSRs. The obvious difference is in capacity and numbers of fibers, receivers, and so on. A 4F OC-n BLSR yields a full OC-n of bidirectional working capacity. A 2F OC-n BLSR yields only half the channels of the OC-n rate as working channels. The other half are set aside for matching protection. On the other hand a 2F BLSR uses half the fibers and transceiver circuit packs, etc. Less obviously, the 4F BLSR has theoretically higher service availability because it can support normal 1:1 APS switching within a span as well as loopback-type BLSR ring action (to follow). For more analysis of the relative merits of 2F and 4F BLSRs, Wu’s book is recommended [9].
4.3.1 BLSR / MS-SPRing Protection Switching

Following a node or span failure, the nodes adjacent to the fault will detect the condition and initiate the ring APS protocol. Each node detecting a fault sends an APS request to the node to which it was connected in the direction of the fault (i.e., the node that is sending it the failed signal). The case of a bi-directional fiber cut between nodes C and D is illustrated in Figure 4. Here, both nodes C and D detect the failed signal. Node C sends D a bridge request with an indication of the nature of the failure (e.g., signal fail), and node D likewise sends C a bridge request. The byte K1 and K2 definitions for this APS communication are shown in Table 2.

Figure 4 Two-fiber BLSR illustration with cable failure

---

4 Fault detection is based on line level measures such as LOS, etc., explained in PMC-Sierra white paper PMC-2030895.
Table 2  Byte K1 and K2 definitions for a SONET BLSR (and SDH MS-SPRing)

<table>
<thead>
<tr>
<th>Byte K1</th>
<th>Byte K2</th>
<th>Bits 1-4</th>
<th>Bits 6-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111</td>
<td>Source node ID is set to the node’s own ID</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1110</td>
<td>Forced Switch [Span] (FS-S)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1101</td>
<td>Forced Switch [Ring] (FS-R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1100</td>
<td>Signal Fail [Span] (SF-S)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1011</td>
<td>Signal Fail [Ring] (SF-R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1010</td>
<td>Signal Degradation [Protection] (SD-P)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1001</td>
<td>Signal Degradation [Span] (SD-S)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>Signal Degradation [Ring] (SD-R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0111</td>
<td>Manual Switch [Span] (MS-S)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0110</td>
<td>Manual Switch [Ring] (MS-R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0101</td>
<td>Wait-To-Restore (WTR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0100</td>
<td>Exerciser [Span] (EXER-S)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0011</td>
<td>Exerciser [Ring] (EXER-R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0010</td>
<td>Reverse Request [Span] (RR-S)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0001</td>
<td>Reverse Request [Ring] (RR-R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0000</td>
<td>No Request (NR)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Destination Node ID is set to the value of the ID of the node for which that K1 byte is destined. The Destination Node ID is always that of an adjacent node (except for default APS bytes).

NOTE – Reverse Request assumes the priority of the bridge request to which it is responding.

When nodes A and B see these APS requests, they allow the APS channel to pass through transparently. When C and D see each other’s requests, they loop back the signals that they were transmitting on the C-D span so that they go around the ring in the opposite direction over the protection channels. This loopback is known as a ring bridge (i.e., a bridge of the affected working channels onto the ring protection channels—analogous to the head-end bridge in APS). At the same time C and D do a ring switch (i.e., the APS tail end transfer function) to substitute the output of the protection channels for the failed working channels. They then update their outgoing K2 bytes to indicate that they are in the bridged and switched state. If only one direction is affected, then only one node will need to perform the ring switch.

In Figure 4 the working and protection channel pair can either be STS-1/STS-Nc (VC-4/VC-4Nc) channels in the same fiber on a two-fiber ring, or they can be separate fibers of a four-fiber ring. In the case of a four-fiber ring, individual span failures between nodes can be protected through a 1:1 type of span switch. Span switch requests are signaled over the affected link between the nodes, which is referred to as the short path. Since it would be possible to support multiple simultaneous span switches around the ring, span switches are given higher priority than ring switches. The existence of a span switch is signaled around the ring (i.e., on the long path) by the switching nodes so that a node needing to request a ring bridge will defer to the span switch and not signal its ring bridge request. The interaction of span and ring switching on four-fiber rings, and the desire for consistent, predictable behavior, have added a large degree of complexity to the BLSR protocol. For long haul carriers, the added protection capabilities of the four-fiber rings have made them worthwhile. For networks covering less geographical area, the relative simplicity of the two-fiber rings has sometimes made them more attractive. In some circumstances, it has even been advantageous to deploy two separate, overlaid two-fiber rings instead of a single four-fiber ring.
A BLSR can also recover much of the ring traffic in the event of either a node failure or even multiple span failures. Consider the case of the node D failure in Figure 5 where D was in intermediate node for a circuit between B and C. The failure is detected by nodes A and C, which will both respond by sending bridge requests to node D in the opposite direction around the ring. When A and C see each other’s bridge requests to D, they know this means that a failure has removed D from the ring. As a result, A and C will then perform ring switches.

Figure 5 Two-fiber BLSR illustration with node failure
4.3.2 Prevention Against Traffic Misconnection

These ring switches due to node failures, however, have changed the connectivity of the ring in a manner that could potentially result in misconnected traffic, as illustrated in Figure 6. In this example, Circuits 1 and 2 use the same channel (time slot) in different portions of the ring. Due to the ring switches, the traffic from A to B is now dropped by D, and the traffic from D to B is now dropped at A. This unacceptable situation was handled in the SONET/SDH standard in the following manner. Each ring node has one database that shows the ring connectivity (i.e., the sequence of node ID numbers around the ring) and a second database that shows the connectivity of each STS-1/STS-\(N_c\) (VC-4/VC-4\(N_c\)) channel on each fiber. (This is an extra step in provisioning that must be performed whenever a path is set up through a BLSR.) When a node sending a ring bridge request receives a ring bridge request from another node, it examines the destination node ID in that request. For a span failure, this destination ID will be its own ID. For the case of a node failure (or multiple failures that have isolated a section of the ring) the destination node ID can be compared to the ring connectivity table to determine which node (or nodes) is missing from the ring. The node will then perform the ring bridge and switch, and simultaneously squelch all STS-1/STS-\(N_c\) (VC-4/VC-4\(N_c\)) channels that were added or dropped at the failed node(s). The other ring node sending a ring bridge request will perform the same functions. This squelching action, which is accomplished by sending AIS-P in these channels, prevents the misconnection.

Figure 6 Potential traffic misconnection problem that a BLSR must handle
Handling potential misconnections becomes more complicated if the ring adds and drops VTs (VC-1/2s). For VT access, it is not desirable to squelch all the VTs in an STS-1 just because a subset of these VTs is added/dropped at the failed node(s). The method adopted for SONET/SDH handles the VT squelching in the following manner. In its initial stages, the protection of the VT accessed ring proceeds in exactly the same manner as an STS-1 accessed ring, including STS squelching. Next, however, once a switching node receives a bridged and switched indication from the other switching node (in byte K2), it will unsquelch those STS-1s that are VT accessed. Once the squelching is removed, it is the responsibility of the nodes dropping those VTs to squelch them at their drop output. To implement this additional VT functionality, the nodes on the ring need additional database information. Each ring node, whether it adds/drops VTs or not, needs to keep track of which STS-1s on the ring are VT accessed so that the node will know which STS-1s to unsquelch if it is a ring switching node. Each node that drops VTs needs to keep track of the other nodes that are the source of the VTs that it drops. Whether a ring node is a ring switching node or an intermediate node that passes through the K1/K2 bytes and protection channels, it will see the K1 ring bridge requests from the two switching nodes. By comparing this pair of crossing K1 bytes, the intermediate nodes can also determine which node(s) have been removed from the ring and use their VT connectivity information to determine which of its dropped VTs originated on the failed node(s). Each drop node will then squelch any of its dropped VTs from the failed node(s) by inserting AIS into the dropped signal.

4.3.3 Extra Traffic and Non-Preemptible Unprotected Traffic

Because the BLSR uses dedicated protection channels, it is possible to use these channels to carry “extra traffic”, also known as standby-line access. This allows low-priority traffic to use the protection channels when they are not needed for protection. When the protection channels are required for the restoration of their associated working channels, the extra traffic is removed (preempted). Carriers typically sell extra traffic bandwidth at a discounted rate, which gives the carrier additional revenue from the otherwise unused bandwidth, and allows cheaper bandwidth for the subscriber with the understanding that the availability is not guaranteed. One additional interesting wrinkle in BLSRs is the ability to support non-preemptible, unprotected traffic (NUT). When a channel that would normally be a working channel is assigned to carry NUT, the corresponding protection channel may or may not also be assigned to carry NUT. When a failure affects the NUT-bearing channels, the NUT will be lost. A failure of the working channels, however, will not affect a NUT bearing channel that would have been a protection channel in non-NUT applications. In other words, a failure on the ring will not result in the traffic on this NUT channel being preempted like a protection channel carrying Extra Traffic would have been. One of the main applications for NUT is to carry traffic that is protected by some other means, such as data traffic protected by a Layer 3 restoration mechanism.

4.4 Ring Protection Processing Challenge

While UPSR implementations can be taxed to process all tributary status information continually, one of the challenges in designing BLSR nodal equipment is that each node must quickly compare the destination node ID of a received ring bridge request with its ring topology databases, and then, if a node is missing from the ring, the switching node needs to squelch the affected channels. In a four-fiber BLSR, it must also evaluate whether the failures should be handled through span or ring switching. The ring bridge loopback and squelch functions are typically performed in the cross-connect circuit of that NE. This squelching-speed challenge for the BLSR comes when a ring failure occurs that removes a node or whole portion of the ring.
4.4.1 TFI-5 Interface

The OIF (Optical Interworking Forum) has recently begun work on a chip-to-chip interface between framer circuits and switch fabric circuits. This interface is called TFI-5 (TDM Fabric to Framer Interface). The TFI-5 interface, which could be confined to a printed circuit board or go over a system backplane, contains overhead channels that can be used to simplify protection implementations. The idea is that the evaluation of the signals (at either line or path level) can be performed by the framer circuits and communicated over the TFI overhead bytes to the switch fabric. A state machine associated with the switch fabric chips can then determine how to reconfigure the switch fabric to implement the required protection switch action. This approach can be used for linear protection, UPSR, or BLSR and allows all the decision processing to take place in the unit that contains the TSI fabric, and distributes the processing load by having the line cards determine the fault or degradation condition. PMC-Sierra’s CHESS III SONET/SDH products implement a superset of the TFI-5 interface and provide extensive integrated hardware processing for linear, UPSR, and BLSR protection switching. This hardware support substantially reduces the real-time protection switching burden on the system software.

4.5 Interconnected Rings

Ring protection can provide a building block for creating larger protected networks by interconnecting rings. For example, a long haul backbone ring can have multiple metro rings subtended from it. Also, a cascade of smaller rings allows more protection capability than a single large ring. A desirable property of ring-based networking in general is that rings are closed subsystems; problems in one ring can be handled in that ring and not require (or cause) actions to be taken in an adjacent ring. The failure of a node associated with the interconnection between the rings is, however, not so clear-cut and needs some special consideration.

In general, it is desirable to have rings interconnect at two geographically distinct locations so that a failure at one interconnection point can be protected by the other interconnection point. Dual-interconnected rings are illustrated in Figures 7 and 8. Because there are two points at which rings are inter-connected, this scheme is commonly called dual-ring-interconnect (DRI) or dual homing (DH). It is also known as the ‘matched node’ arrangement by some vendors. A drop-and-continue function is typically used in which the interconnecting traffic is dropped at one of the interconnection nodes and then continues around the ring to also be dropped at the other interconnection node. UPSRs and BLSRs differ in how they select the best interconnection signal.

5 “Dual homing” is a BellCore term while T1 and ITU use dual-ring interconnect to refer to the same architecture for ring-to-ring interconnection. Note that DH also refers to the survivable PSTN architecture where a remote access concentrator or small CO connects to two hub COs for protection against failure of either hub CO.
4.5.1 Interconnection with UPSRs

In a UPSR (Figure 7), each interconnection node performs a path switch function on the traffic that it sends to the other ring so that the best path is chosen by the same principles discussed above for path switching. A UPSR receiving traffic from another ring behaves somewhat differently than a normal UPSR. Rather than an interconnection node adding the traffic in both directions around the ring, one interconnection node adds the traffic in one direction and the other adds it in the opposite direction. The destination drop node then performs the same type of path switch function as a normal UPSR drop node.

Figure 7 Dual Node Interconnection of UPSRs

4.5.2 Interconnection with BLSRs

For BLSRs (Figure 8), the line switching eliminates the need to perform path switching on the traffic dropped from the feeding ring to the receiving ring. The receiving ring, however, must choose the best signal coming from the two interconnection nodes. Looking at Figure 8, consider the example of a failure of interconnection node C in feeding ring J. This failure would mean that node M in receiving ring K is receiving good traffic from its connection to node D and failed traffic from its connection on the ring to node N. The selection process at node M is referred to a service selection and is essentially the same as path switching except that the selection is made on add/pass-through signals rather than a dropped signal. Figure 8 shows an example of opposite routing in which the signals in one ring are routed around the “opposite” side of that ring as the signals are routed in the other ring. Same side routing, which is also allowed, would place the service selection function in node N in Figure 8. Other variations on BLSR interconnection involve the methods used for carrying the drop-and-continue traffic between the interconnections nodes on a ring. For example, the continue portion of the drop-and-continue traffic can be placed on the protection channel so that it doesn’t reduce the ring’s available working bandwidth. UPSRs and BLSRs can be interconnected to each other using a combination of the interconnection techniques of Figure 7 and Figure 8.
4.5.3 Interconnection with Intermediate Multiplexing and Payload Defect Indication (PDI)

Two general techniques are shown in Figure 9 for connecting the interconnection nodes of one ring with those of the other. For either technique, the physical connection can be a SONET/SDH or asynchronous/PDH signal. In the first interconnection technique, the logical connections between the nodes in the rings are made directly at the same level in both rings (e.g., VT1.5, STS-1, DS1, E1, or DS3). In the second interconnection method (Figure 9 (b), there is an intermediate NE between the interconnecting ring nodes that performs a multiplex function. This multiplex function is typically used to add traffic from other locations. For example, if ring W was a VT-accessed ring (e.g., from a metro network) and ring Z was an STS accessed ring (e.g., from a long haul network), the intermediate multiplexer or cross-connect would take VTs from ring W and VTs from other sources and multiplex them into a STS-1 that it hands off to ring Z.

Figure 9  Ring interconnection examples

a) Direct dual-fed interconnection

b) Dual-fed interconnection through an intermediate multiplexer or crossconnect
The intermediate NE thus allows more efficient fill of the STS-1s in ring Z. (Clearly, the same external traffic would need to be added at both intermediate multiplexing points. This traffic could come, for example from a different ring.) When the destination drop node is in a STS accessed UPSR, the example in Figure 9 (b) leads to an interesting problem. A failure of one of ring W's interconnecting nodes would mean that the STS-1 formed by the intermediate NE would comprise failed VTs in the time slots associated with ring W traffic. Meanwhile, the STS-1 handed to ring Z through the other interconnection node would be comprised of good VTs from ring W due to ring W's protection switching around the failed node. Since ring Z does not process VTs, it would have no way of knowing that the constituent VTs of one STS-1 are good while those of the alternative STS-1 are failed. Hence, the drop path switch in UPSR Z could drop a STS-1 with bad payload when a STS-1 with good payload was available. It is clearly not desirable to require a STS drop node to have the circuitry to examine every constituent VT Path. The alternative was to define a Payload Defect Indicator (PDI) that is generated and used in the following manner. An intermediate NE is required to keep track of how many incoming tributary signals are bad. This number is encoded into the SONET/SDH C2 byte (see Table 4 of PMC white paper PMC-2030895). When the STS UPSR drop node is receiving two good STS-1s, it examines the C2 bytes of each to determine the relative health of their payloads. The STS-1 containing the lowest number of failed constituent tributaries is then chosen as the dropped signal.

4.5.4 Other Ring Interconnection Considerations

Two points warrant closing emphasis, however. First, keep in mind that dual-ring interconnect is a per-path decision to treat the signal this way. The two rings involved need to topologically share two so-called “matched nodes” to enable the possibility of DRI, but DRI itself is not a whole-ring to whole-ring relationship or concept. In other words it is not a pair of rings that are DRI connected as a whole, or not, but rather each path transiting ring to ring is or is not set up with DRI. The decision can be made to provision each individual path with a single geographical ring-to-ring transition point, or to set up a drop-and-continue arrangement for DRI. Secondly, if readers imagine that DRI setups, especially the drop-and-continue aspect, tend to rather quickly exhaust the span between the matched nodes, they are quite right. The premature exhaustion of provisionable capacity between the matched nodes is one of the issues with DRI. For this reason, variations such as “drop and continue on protection” have been proposed where protection capacity is used for the continue signals. If an intra-ring switch occurs, the continue signal may be temporarily bumped off. Otherwise if an inter-ring transition failure occurs, the continue signal is intact, but without consuming working capacity. There is a theoretical availability reduction, but in practice this may be less important than extending the provisioning life of the rings. It is also possible to move the drop signal over to the other ring and provide its continue function in the other ring itself but again this is not without some loss of theoretical availability.

6 The “cross-office” (in-building) wiring to go from an ADM drop on ring X over to the add interface on an ADM for ring Y, will often be via an OC-48 or higher rate 1+1 or 1:N APS protected transmission system with “short-reach” optics. So DRI really is a measure to protect against the possible loss of an entire building (presumably via fire, power loss, sabotage), not primarily just the cross-office cabling.
Further discussion of strategies for dual-ring interconnection, including detailed availability analysis and resource implications of a variety of ring interconnect strategies, can be found in [10] [11] [12] [13] [14]. A good general reference for more details on rings and APS technology is Wu’s book [9].

5 Partial Protection Using the Link Capacity Adjustment Scheme (LCAS)

LCAS [6] works with virtually concatenated channels, as described in PMC-Sierra white paper PMC-2030895, to create and manage physically diverse routings for component signals of concatenated SONET pipes. For data-centric applications, this provides a new service restoration option. A virtually concatenated channel is one in which a number of paths are combined to form a larger channel such that the individual constituent paths (i.e., the members of the virtually concatenated group (VCG)) can take different routes through the network. The virtual concatenation sink node realigns the members so that the payload data can be extracted from the VCG channel. A carrier can provision the VCG such that the members are intentionally routed along diverse routes. When a failure on one of the routes removes some of the members, LCAS signaling from the VCG sink tells the VCG source which members have failed. The LCAS source then stops using the failed members and falls back to using only the healthy members. The service is then restored at the lower bandwidth.

This restoration is illustrated in Figure 10. Figure 10 (a) shows three physically diverse LCAS paths along with the allocation of STS-1 constituent flows to the three paths of the LCAS group. In the pre-failure state, the service seen by the end nodes is equivalent to an STS-5c (hence the five SQ numbers). Fully disjoint routes will obviously be a common preference where possible, but in general as shown, efficiencies in bandwidth-use can be achieved by defining multiple diverse routes that are not mutually disjoint. (For example here, the downward diversion that SQ0 takes for part of its route, before re-joining SQ1, could be warranted if the link that SQ0 takes alone is restricted in available bandwidth.)

In Figure 10 (b) a failure occurs on a span that is common to two of the three paths forming the LCAS VCG. At this point, two options exist. The failed members can be logically removed from the VCG through a provisioning configuration change, or the VCG can continue to exist in its initial logical form but using only the bandwidth of the surviving path until the failure is cleared. In the second scenario, when the failure clears, the sink will see signaling indications that the source is still not placing data in these members (although those paths are now noted to be working again). When the source receives an OK message for the restored members, it will resume placing data in them.

See Chapter 4 for the signaling details.
Although this is not protection or restoration of the pre-failure state as seen by the application flows, in the sense of the other protection techniques in this chapter, this form of response to failure can be highly attractive for cost-sensitive data applications. LCAS allows use of unprotected links with a fallback to lower bandwidth in the event of a failure, rather than suffering a complete loss of the connection or having to pay for full protection. Thus LCAS gives a form of graceful degradation rather than a complete and sudden outage that could arise from failure of a single unprotected service path. Also note, however, that nothing prevents one or more of the constituent LCAS member paths from itself being routed with conventional protection. This would allow one to design services that exploit extra bandwidth when available (by adding LCAS members), with fall back upon failure but only to a certain point—strongly ‘drawing the line’ at a minimum LCAS capacity represented by the protected-service group members.

**Figure 10** LCAS fallback to reduced bandwidth due to failure affecting two out of three members of the VCG

*a) Initial, provisioned state of the VCG with diverse routing of members*

*b) State of the VCG in the response to the failure on a link affecting two members*
6 Conclusions

Fiber optic transmission allowed a vastly increased amount of traffic over each span. For example, a 9.95328 Gbit/s OC-192 signal can carry up to 129,024 voice channels. Clearly, it would be unacceptable to allow single failure points in the network to disrupt such a large number of subscribers. Also, as we have grown to rely ever more on our communications for conducting business and such critical functions as coordination of air traffic control, network failures that interrupt service for a noticeable period of time are unacceptable. This situation was addressed by SONET/SDH at the outset by incorporating protection switching control channels into the signal overhead. While SONET/SDH began with the relatively simple point-to-point line protection scheme, it has proved to be very extensible in terms of topologies (e.g., rings) and flexibility (e.g., LCAS service restoration). The 1+1, 1:1, UPSR, and BLSR technologies have each proven to have an important role in the network. LCAS-based restoration shows promise for increasing the cost effectiveness of many packet-based services.

The APS technologies described in this white paper are all capable (in fact, required) to complete circuit restoration within 50 ms of failure detection. This speed requirement is typically the most difficult real-time task for network elements, especially for ring add/drop multiplexers (ADMs). The ring ADM protection tasks require efficient communication of the fault detection to the protection control sub-system and efficient communication to the switching or cross-connect fabric in order to implement the protection switch.

PMC-Sierra's SONET/SDH components have been designed with these real-time constraints in mind. The CHESS III devices implement a superset of the TFI-5 protocol on the interface between the framers (which detect the faults) and the cross-connect fabrics that includes a protection messaging channel to quickly and efficiently communicate the type and relative severity of a failure. Along with the messaging channel, integrated processing functions within the framers and cross-connect devices can be configured to execute protection switching events automatically without direct software intervention. The advanced APS implementation of the CHESS III devices allows them to meet or exceed required switchover times and offload the significant software processing required for APS functions in high capacity systems.
7 References

[1] American National Standards Institute, T1.105-2001 - Synchronous Optical Network (SONET) Basic Description including Multiplex Structure, Rates, and Formats


