

Protection and Restoration in DWDM Networks: Recent Developments and Issues

Sudheer Dharanikota and Raj Jain¹
Nayna Networks, Inc
481 Sycamore Drive
Milpitas, CA 95035
Sudheer@nayna.com, raj@nayna.com

1 Abstract

Protection using rings has been the common method of reducing downtime on SONET/SDH based telecommunication networks. With the trend towards IP over DWDM, there is a need for IP based protection and restoration mechanisms. In this paper, the mechanisms being developed at several standards organizations including OIF, IETF, and ITU are described. These mechanisms are in some sense more powerful than the previous SONET/SDH based mechanisms and protect against not only link and node failures but also against domain failures. Here domain refers to an entire region of the network. In particular, Shared Risk Group (SRG) concepts being developed at IETF and OIF is explained.

2 Introduction

In telecommunications networks, protection and restoration refers to mechanisms used to minimize the downtime due to failures. The difference between protection mechanisms and restoration mechanisms has been a matter of debate at various standards bodies. To avoid this debate, we address them as recovery mechanisms in this paper. Telecommunications network architectures consist of three component planes: data, control and management. Data plane consists of components and protocols required to transmit data on a given path. The path itself may be determined manually or automatically using an intelligent control plane. The management plane helps monitor and manage the faults, configuration, accounting, performance and security (FCAPS) of the network. Although, these three planes have distinct functions but may or may not be physically separate. For example, a SONET network may be controlled by an IP-based control plane. The IP messages may be sent over the same SONET network or may be sent over a separate data communication network (DCN). Failures can occur at any of these three planes, but *in this paper we are only interested in the data plane failures and their recovery with the assistance of the control plane.*

As stated above, one component of management plane is "configuration" which involves planning the network topology. Current and projected future traffic matrix can be used to determine the most effective topology. Recovery considerations in case of faults may further require changes or additions to the topology. Once the network has been configured and installed, as requests for connection come in, the paths may be provisioned manually or using control plane. The latest development in this area is the development of Generalized Multi-Protocol Label Switching (GMPLS) [gmpls-ospf, gmpls-isis], which allows provisioning a path over networks consisting of multiple technologies including fiber switching, wavelength switching, time division multiplexing, and packet switching. GMPLS includes mechanisms for requesting a connection, determining the optimal path through the network, and also determining or setting up resources that will help in recovering from faults. Several of these mechanisms are described further later in this paper.

A telecommunications networks may consist of several networks owned by different carriers as shown in **Figure 1**. Recovery resources have to be provisioned in each carrier's network as well as between carrier networks. In other words, both intra-carrier and inter-carrier recovery issues are important. *In this document we focus on the intra-carrier recovery mechanisms.* A carrier network may be further broken into smaller domains based on administrative reasons, vendor separation or other reasons as shown in **Figure 1**. The nodes on the edges of these domains, e.g., B and C or D and C in the above figure are called Domain border nodes.

¹ Corresponding Author

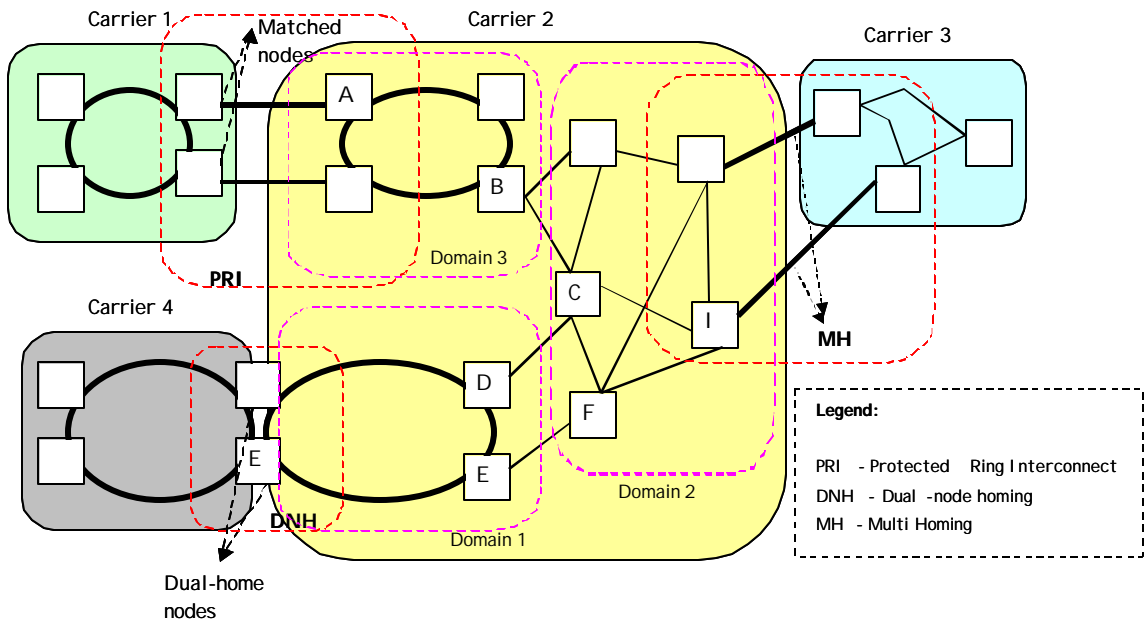


Figure 1 Different inter and intra-carrier interconnecting scenarios

The Carrier working group at Optical Interworking Forum (OIF), has developed a set of requirement guidelines for transport networks. The intra-carrier recovery requirements include [oif-carrier-p&r-reqs]: support for client signal independence, priority based connection recovery, single failure (at a minimum) recovery, bulk connections recovery; support of intra-domain, inter-link, inter-domain and end-to-end recovery mechanisms, support for multi-layer recovery, support for low priority traffic occupancy over restoration resources etc.

3 Recovery Related Classifications

3.1 Fault Classification

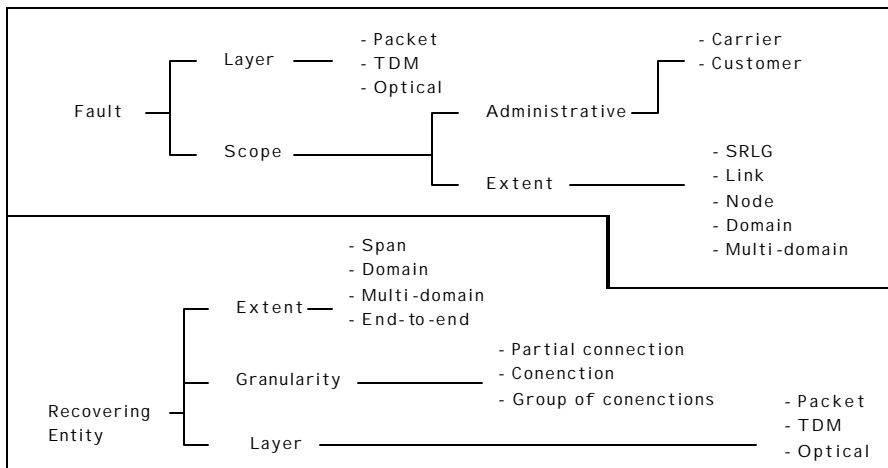


Figure 2 Fault and recovering entity classifications in a typical transport network

The main reason for protection and restoration is to help quickly recover connection(s) in case of failures. Failures and the affected entities can be classified as shown in Figure 2. Typically, data packets are transported over TDM connections through a network of fibers. In this case, the network is said to have three layers: Packet, TDM, and Optical. A failure may be relevant to a particular layer due to detection and recovery actions that can be performed on such a failure. Failures are also classified by the scope of recovery. This scope can be administrative scope and/or the extent of failure. The administrative scope is used to determine the boundary of recovery, that is, if the recovery is to be performed in a single carrier domain or if it is to be propagated to another network. The extent of recovery could be a span, a node, a domain, multiple domains, or end-to-end. The layer, which performs the recovery action, could be the packet, optical or TDM layer. The entity that is recovered could be a partial connection, a connection or a group of connections (such as link recovery). The connections that are recovered could be based on their restoration priority.

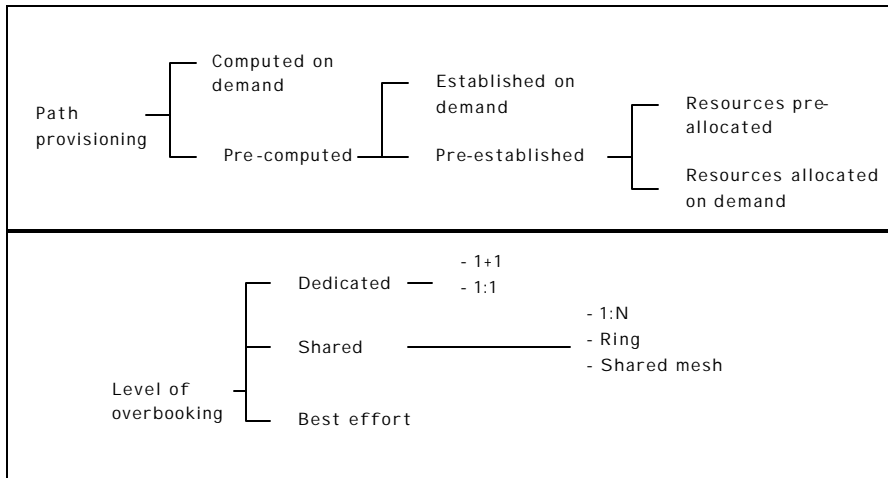


Figure 3 Path provisioning and over booking classification in a typical transport network

Proper path provisioning helps in alleviating the effect of the failures. As an example, one may compute primary and secondary paths to protect a connection from any single link or node failure. In this case, the two paths will not share any link or nodes. Not only will they not share any link, the two connection will not have any links that can fail simultaneously. For example, if two links pass through a single conduit, the conduit failure could bring both the links down. The group of links that can fail simultaneously is called Shared Risk Link Group (SRLG).

The path provisioning can be categorized, as shown in Figure 3, based on when the secondary path is computed (pre-computed or computed on demand), when the secondary path is established (pre-established or established on demand) and when the resources are allocated to the secondary path (pre-allocated or allocated on demand). Note that these different options provide different connection restoration times. There are many mechanisms available in the literature to overbook the resources (i.e., compute/establish/allocate the secondary resources) to cope with the failures. This overbooking can be done per connection, per link (also known as span protection) or per domain (as in ring topologies). In all these cases the level of overbooking, as shown in Figure 3, can be classified as dedicated (such as 1:1, 1+1), shared (M: N, Ring, Shared mesh) or best effort (recovered only if the resources are available). Under shared restoration one may support preemptable (preempt low priority connections in case of resource contention) traffic and non-preemptable traffic.

3.2 Recovery Process Related Classification

A recovery may involve many entities as shown in Figure 4. They are: the detecting entity that detects a failure or group of failures, the deciding entity that makes the recovery decision, and the reporting entity that correlates the failures, groups them if necessary, and reports them the deciding entity. Communication among these entities can be through the data channel or through a separate channel. These two types of communications are known as in-band and out-of-band communication, respectively. *In the following sections we only consider the requirements instead of the mechanisms to be independent of in-band and out-of-band discussion.*

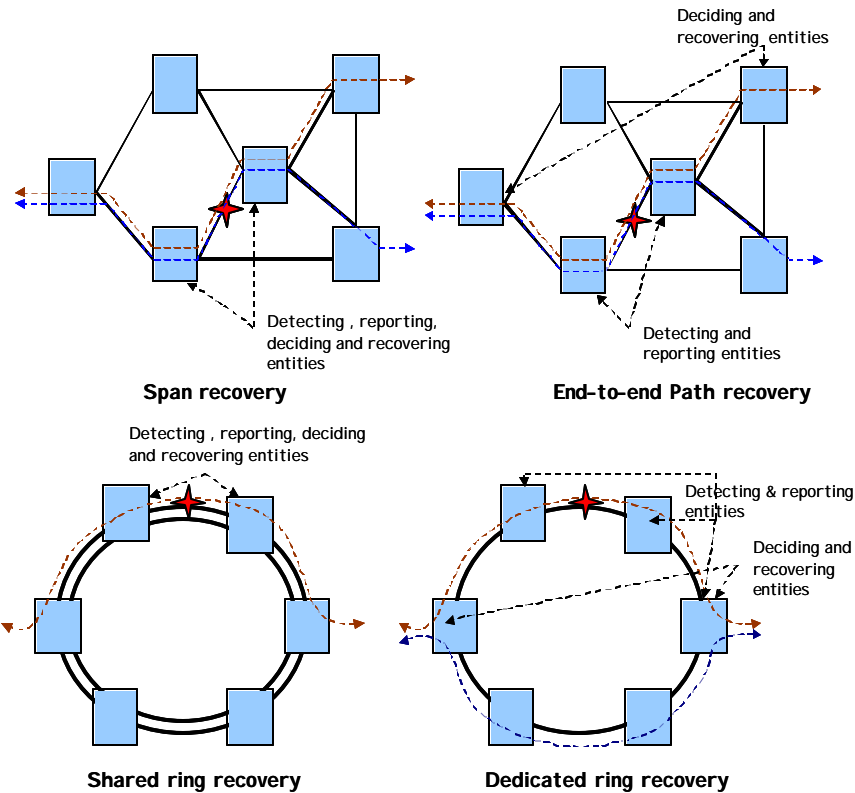


Figure 4 Different examples to illustrate the entities involved in recover ring a link or SRLG failure

4 Applications

We focus on connection recovery in a single layer. Usually, each sub-network or domain will provide its own recovery functions, such as SONET ring self-healing protection, shared mesh restoration, or dynamic restoration. If a connection travels multiple sub-networks or domains, then the recovery may be either domain-by-domain recovery or end-to-end recovery. Several different such applications are described below.

Metro/core application with domain recovery: Figure 5a shows a simple application of metro-core-metro connection. The solid line indicates the service path while the dotted line indicates the recovery path. The service path is partitioned into 5 segments and each segment may be recovered individually. In this scenario, we assume that each metro has its own protection/restoration scheme and the core is treated as a single domain with a standardized protection/restoration scheme. A path needs to be protected from many types of faults that can occur in such a scenario. In the figure, we present two such faults. Fault A, which is on the link between the domain-edge nodes, can disrupt both the working and protecting paths. Connections can be protected from Fault A by allocating protection resources (dedicated or shared) on a parallel link. By this allocation one can open up different recovery mechanisms using these additional resources, which in turn leads to different recovery processes. Note that we may not recover from all types of faults (for example a node failure) in such a scenario, which could lead to a crank-back to the source for an alternative path. Also note that if there is a contention for the protection resources, a priority-based scheme should be used to resolve the contention. On the other hand, Fault B in the core domain can be protected using a path that does not share the risks of the working path. Two links that can fail simultaneously with a single fault are said to belong to a single "Shared Risk Link Group" or SRLG. Two paths that consist of links that do not share any risk and cannot fail simultaneously by a single failure are called SRLG-diverse paths.

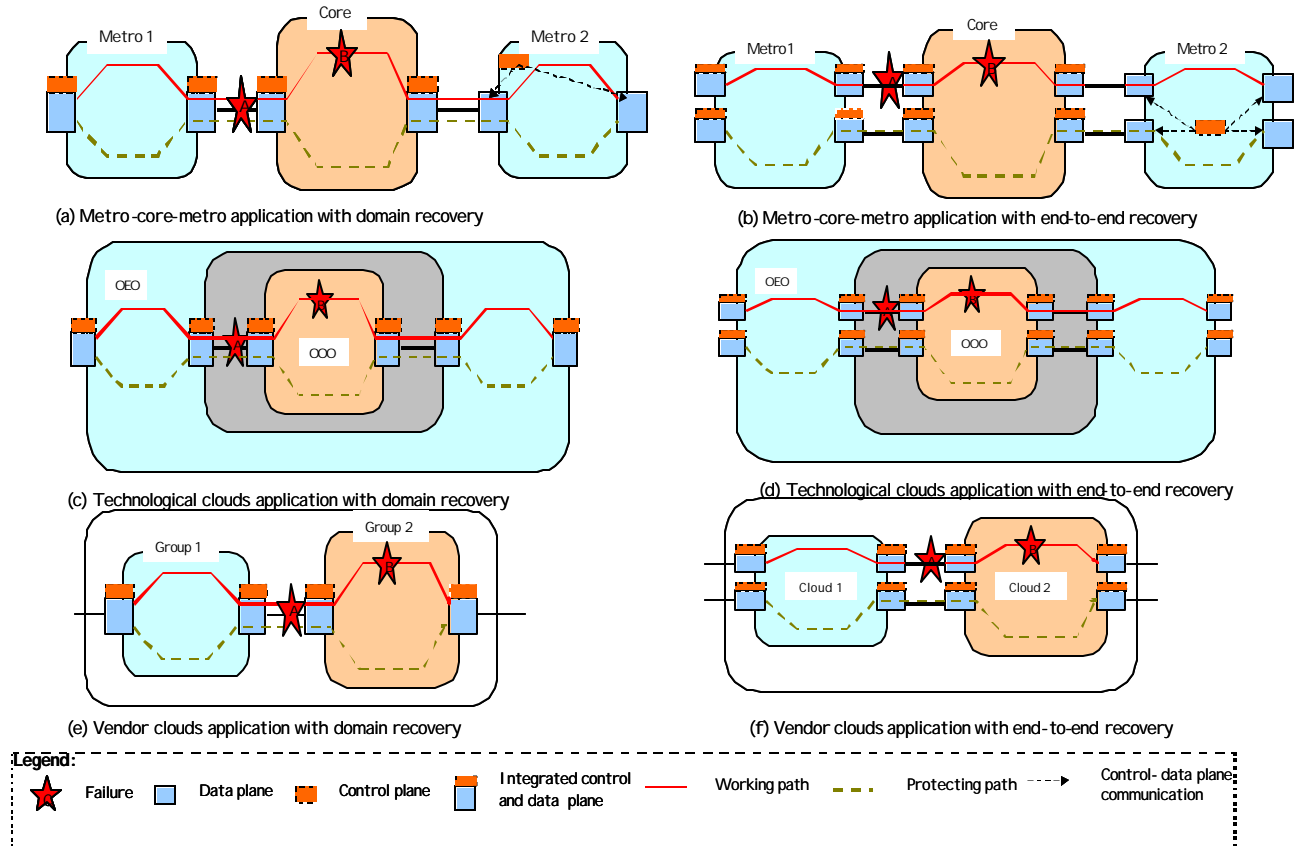


Figure 5 Examples of Recovery Applications

Metro/core application with end-to-end recovery: When there are no trust and policy restrictions between the domains, one can compute a strict end-to-end path crossing multiple domains. Figure 5b shows a simple example of end-to-end recovery of a connection crossing three metro-core-metro sub-networks. In this scenario, the recovery path should be completely node-diverse an SRLG-diverse from the service path such that the connection can be recovered from any single node or link failure. Unlike in the previous example, an end-to-end path recovery can protect the connection from most of the failures. The disadvantage here is the time it takes to recover from a failure. Note that in such a scenario many recovery mechanisms can be employed with multiple levels of overbooking at different places in the network. Contention for protecting resources occurs in this scenario also.

Domain of transparency application with segment recovery: With the technology improvements of optical networks, some carriers may build a high-bandwidth long-haul transmission network over existing sub-networks to reduce cost. For example, an all-optical domain (OOO) may be added to the optical-electronic-optical (OEO) domains. A connection crossing the core may travel an OEO domain, enter the OOO domain, and then return to the OEO domain again. Figure 5c shows a connection traveling overlay-domain with segment recovery application. The connection is partitioned into 5 segments and each segment is protected individually. In this application, the nodes connecting the internal subnetworks are not protected. These nodes form the so-called internal network-to-network interface (I-NNI).

Domain of transparency application with end-to-end recovery: Figure 5d shows an example of end-to-end recovery with overlay-domain scenario. Again, the solid line stands for the service path and dotted line for recovery path. In this application, the two paths can be completely node/SRLG disjoint. However, this application may impact the architecture configuration and routing/signaling protocols.

Table 1 Fault/degradation versus the mechanisms of reaction by OXC and DWDMs

Location	Degradation Or Fault	Actions		
		L-DWDM	PXC	U-DWDM
A	Fault	D, R (Down)		
B	Fault	M, G, R		
C	Fault		M, R	M, G, R
D	Fault		M, R	M, G, R
E	Fault			R
F	Fault			R

The LMP-DWDM protocol between the DWDM and the PXC supports the following proposed features: Event driven and polling based fault and performance reporting with thresholds; Recording the history of the monitored parameters; customized error reporting; specifying and negotiating the parameter set to be monitored; and negotiating the threshold values. Additional facilities may be added in future proposals to add the following features: control channel management, link property correlation, connectivity verification, negotiating the loss of light (LOL) behavior, group reporting to reduce the overhead, and path (optical trail) tracing mechanism

6 Signaling

The generic requirements for a signaling protocol to support recovery mechanisms are: they must support fault notification constructs to domain edge nodes and to the source, and must support carrying information related to different link recovery (such as dedicated and shared) mechanisms, and must support different preemption levels and usage of protecting resources in normal cases. They may support fault isolation constructs.

6.1 Communication Between Detecting and Reporting Entities

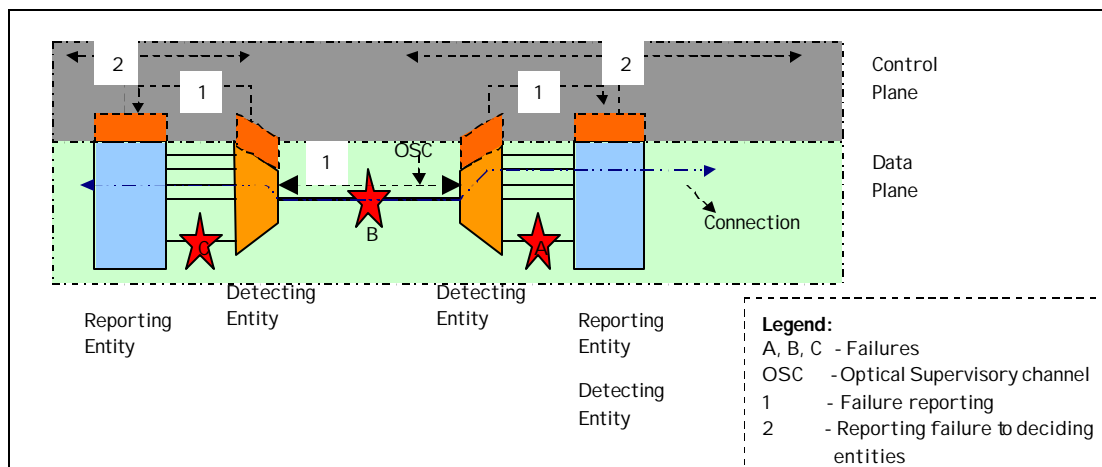


Figure 7 An example configuration to illustrate different communication mechanisms between detecting and reporting entities.

The following are the cases considered to determine the requirements on the communication between the detecting and the reporting entities:

- i. Both the detecting and reporting entities are in the same box (e.g., SONET equipment, Opaque cross-connects, Some cases of transparent cross-connects etc.). This is the case for failure A as shown in Figure 7.
- ii. Detecting and reporting entities are separate but have in-band communication between them (e.g., SONET APS, OXC's LOS, etc.). This is the case for failure B in Figure 7.
- iii. Detecting and reporting entities are separate but have out-of-band communication between them (e.g., OXC, PXC's LOL). This is the case for failure C in Figure 7.

General requirements include building a relation between the view of the failing entities to the connections both from the detecting and the reporting entity point-of-view. Detecting entity should know the reporting entity unless the failure is automatically detected by the reporting entity too. Detecting entity should be able to group (or correlate) as many errors as possible before reporting. After the failure, the detecting entity should communicate the failure(s) to the reporting entity.

6.2 Communication Between Reporting and Deciding Entities

The following are the cases considered to determine the requirements on the communication between the reporting and the deciding entities:

- i. Both the reporting and the deciding entities are the same (e.g., span protection in both ring and mesh networks). This is the case represented by 2.1 in Figure 8. For example, this can normally be performed by SONET APS mechanisms (in-band) or using out-of-band control protocols such as LMP, RSVP etc.
- ii. The reporting and the deciding entities are not the same but have in-band communication between them. This is the case represented by 2.2 in Figure 8. For example, this can be achieved through SONET APS like mechanisms [SONET].
- iii. The reporting and the deciding entities are not the same but have out-of-band communication between them (end-to-end path or MPLS fast restoration like mechanisms). This is the case represented by 2.3 in Figure 8. For example, this can be achieved through newer APS like mechanisms [G.841].

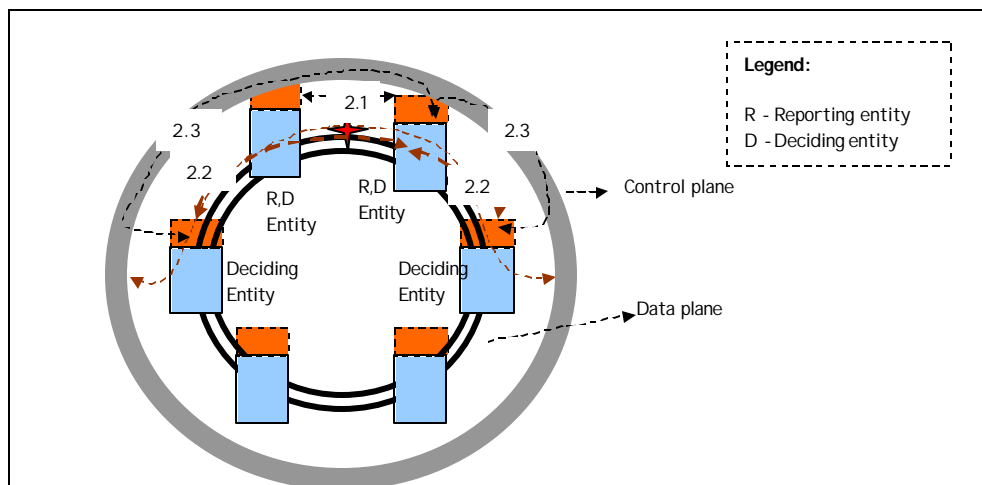


Figure 8 An example configuration to illustrate the communication mechanisms between the reporting and deciding entities.

General requirements for such support are that the reporting entities should know the deciding entities (to whom to report), the deciding entity should have a relation between the connection(s) and the failure recovery action to be performed. The reporting entity should indicate a group failure whenever possible. The communication requirements

include informing the failures (individual or group), and the connection status and the type of failure, if available, to both deciding entities. Note that in case of a span recovery these two deciding entities could be neighbors. In-band or out-of-band communication should be present between the neighbors to report failure(s). If the deciding entities are multiple hops away then the failure can be communicated using a directed message to them. The deciding entity could be a centralized management system, in which case a communication channel should be present between the reporting entity and the deciding entity.

6.3 Communication between Deciding and Recovering Entities

The following are the cases considered to determine the requirements on the communication between the deciding and recovering entities:

- i. Both the deciding and recovering entities are the same (e.g., span protection, wrap around decision in ring networks). This is the case if A and C, and H and D are the same in Figure 9.
- ii. The connection end-points are the recovering entities (e.g., end-to-end path recovery in mesh networks). This is the case when A and B, and E and H are the same in Figure 9.
- iii. The domain end-points are the recovering entities e.g., ring, protection against failure of segments (B-C, C-D, D-E). This is the case when B and E as shown in Figure 9 (case 3.1) perform the recovery.
- iv. Intermediate points are the recovering entities (e.g., shared mesh restoration). This is the case when B, F, G, E, as shown in Figure 9 (cases 3.1, 3.2 together), participate in the recovery.

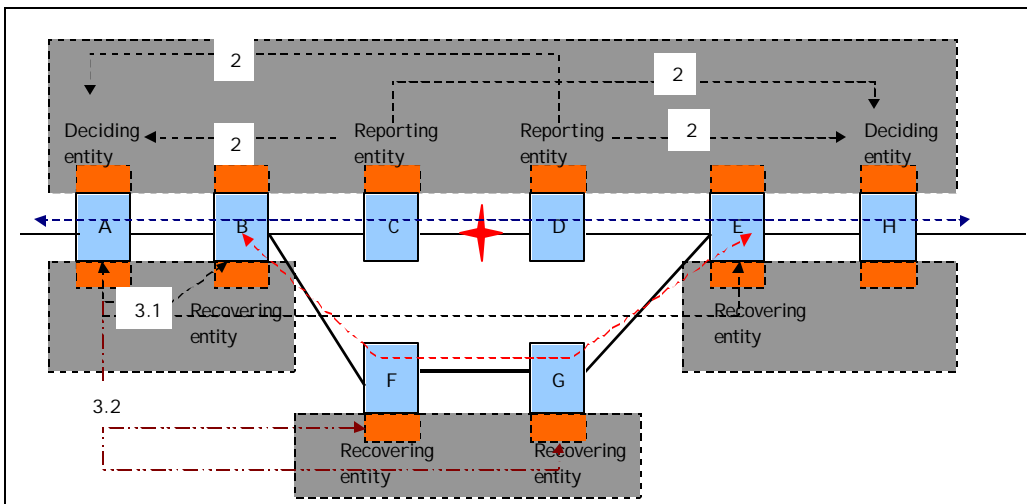


Figure 9 An example configuration to illustrate a communication mechanism between deciding and recovering entities.

General requirements are that the deciding entity should know the recovering entities identification, the deciding (or the recovering) entity should perform contention resolution for the recovering resources and the failed connections should be recovered in the order of their priority. The communication requirements are that deciding entities should communicate to the recovering entities about the actions to be performed. Both the deciding entities and the recovering entities should co-ordinate the switchover operation.

7 Routing

Routing protocols in transport networks are used to communicate the resource properties, which in turn can be used in computing the diverse paths with bounded risk of failures. These resources for an intra-domain case are links or nodes. But for inter-domain case, these are inter-domain links, border nodes, and domains themselves. A diverse path computation algorithm such as modified Dijkstra's algorithm is used to compute these diverse paths. Such an algorithm

takes inputs from three databases namely, *topology*, *traffic engineering (TE)*, and *existing path databases* to provide responses to new path requests, as shown in **Figure 10**. The *topology* and *TE databases* are managed through routing protocols such as OSPF, ISIS or by querying the network management system. The *topology database* contains the nodes, links, and their interconnection. The *TE database* contains the properties or capabilities of the network resources. The *existing path database* contains the path information, such as the nodes and links traversed by various paths in the network. With the above inputs, a request to find diverse paths between a given source and destination pair with given constraints on path selection is processed by the algorithm. The computed path could be a complete enumeration of all intermediate nodes or a partial list of key intermediate nodes between the source and destination pairs. The preceding two options are called strict explicit path and loose explicit path, respectively. Once the response is accepted, the computed path is recorded in the existing path database.

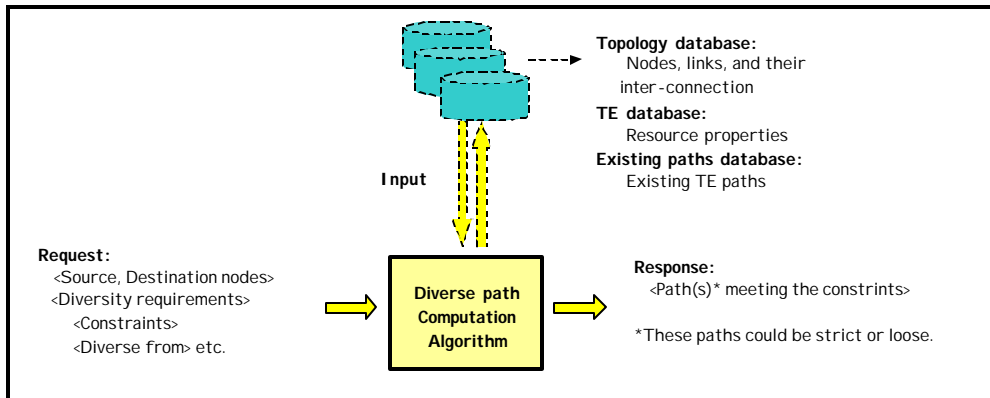


Figure 10 Different interfaces to a diverse path computation algorithm

The diversity requirements of carrier transport networks have some differences from those of packet (Layer 3 and Layer 2 switching) networks. Transport networks inherently provide elaborate protection and restoration mechanisms. These networks are not always structured in mesh topologies as assumed by the packet networks. Transport networks contain multiple sub-layers unlike in packet networks. This leads to different strategies for protection and restoration. At present, in the packet networks only interfaces (or links in some sense) have capability assignment, whereas in the transport networks links, nodes and domains have capabilities. Therefore, the path computation in transport networks allows more elaborate inclusive and exclusive constraints. Also, since the transport networks are grouped differently than packet networks, the path computation mechanisms may not have the complete information about the topology to compute both loose explicit path and strict explicit path.

Risk assessment is defined as the evaluation of the potential risk associated with the inclusion of a given resource in a given path. For Example, consider the following client requests to the optical network:

- Request a persistent connection with 99.999% (widely known five 9's) availability or equivalently a downtime of less than 5 minutes per year or
- Request a higher protection for a portion of the traffic (at the expense of paying a higher charge) compared to other low-priority traffic.

Such requirements will be translated into constraints in path computation. Such constraints can be grouped into path selection constraints and path characterization constraints. The *path selection constraints* typically dictate which physical path should be taken to achieve the client's availability requirements. These requirements are typically the logical and physical diversity. The *path characterization constraints* typically dictate the protection mechanisms as requested by the client. This can be achieved in the form of optical rings, mesh protection mechanisms, etc. These constraints can be satisfied using the link, node, and domain capabilities as discussed in the previous section on diversity. The components that need formalization in this example are specifying the user requirements, translating the requirements into path computation constraints, configuring the network in a way that helps in assessing its features (such as the availability), propagating the information, and finally using information in path computation.

The generic requirements of the protocols are that they should import the recovery capabilities of links between the domains, propagate the recovery capabilities of the other domains that are accessible through a border node (Such a mechanism will help in selecting proper entry point, and they should import some of the protection capabilities of the other domains based on certain policies.

For segment-by-segment recovery scheme, the two border nodes connecting two sub-networks are required to discover the link-based recovery capability for the peer links. If the peer link fails, these two nodes are required to perform the same link-based recovery scheme. For end-to-end recovery scheme, the routing requirements for provisioning are not sufficient for connection recovery. Provisioning needs network topology and resource information to select a feasible and efficient route for a connect request at the first node. End-to-end recovery requires that the first node is able to select a SRG disjoint path from the service path.

8 Conclusions

In this paper, we have provided taxonomy of faults and recovery mechanisms. Faults can be classified based on the layer and scope. Recovering entity can similarly be classified by extent, granularity and layer. Protection paths may be pre-computed or provisioned on demand. In either case, the protection path may be dedicated, shared, or best effort. Several applications including metro-core with domain/end-to-end recovery, domains of transparencies with segment/end-to-end recovery, and inter-vendor clouds with segment/end-to-end recovery were described. Recovery processes require three kinds of entities: fault-detecting entities, fault reporting entities, and deciding entities. These entities may or may not reside in the same box. If they are on different boxes, then protocols are required for communication between these entities. One example of such a protocol, which has been developed recently, is LMP-DWDM, which allows fault communication between DWDM and photonic cross connects. Routing protocols, such as OSPF and IS-IS are being modified to allow easy computation of protection paths.

9 References

- [oif-carrier-p&r-reqs] oif2002.050.03 - Carrier Requirements for Restoration over NNI
- [NNIP&R] oif2001.507.01 - NNI Protection and Restoration requirements
- [P&Rreq] oif2001.367.02 - Multi-layer protection and restoration requirements
- [IDR-SRG] oif2001.227.02 - Inter domain routing with Shared Risk Groups
- [Faults] oif2001.055.00 - Detecting and correlating external path-related faults and degradations
- [G.841] ITU-T Recommendation G.841, "Types and Characteristics of SDH Network Protection Architectures," July 1995.
- [ANSI-T1.105] "Synchronous Optical Network (SONET): Basic Description Including Multiplex Structure, Rates, and Formats," ANSI T1.105, 2000.
- [LMP] J. P. Lang, et al., "Link Management Protocol (LMP)," IETF working group document, draft-ietf-mpls-lmp-02.txt.
- [LMP-DWDM] A. Fredette, et al., "Link Management Protocol (LMP) for WDM Transmission Systems," draft-fredette-lmp-wdm-01.txt, an IETF work in progress.
- [IETF-SIG] B. Rajagopalan et al., Signaling for Protection and Restoration in Optical Networks, IETF work in progress, November 2001.
- [IETF-REST] G. Li et al., RSVP-TE Extensions For Shared-Mesh Restoration in Transport Networks, IETF Internet Draft, work in progress, July 2001.

[gmpls-ospf] K. Kompella et al., "OSPF Extensions in Support of Generalized MPLS," draft-kompella-ospf-gmpls-extensions-01.txt, work in progress.

[gmpls-isis] Kireeti Kompella et al., "IS-IS Extensions in Support of Generalized MPLS," draft-ietf-isis-gmpls-extensions-02.txt, work in progress.

10 Authors biography

Sudheer Dharanikota: Sudheer Dharanikota obtained his Master of Engineering from Indian Institute of Science (IISc) in 1990 and Ph.D. from Old Dominion University in 1997. He worked as a scientific officer in ERNET at IISc for two years on many networking technologies. After his Ph.D., he worked at Racal Datacom as a Manager of Routing, Bridging and Frame Relay compression groups from 1996-1997. Then he worked at Alcatel USA as a manager in many data products, including RCP 7770 - a 640 Gbps core router, from 1997-2000. He is also a research associate professor at Old Dominion University. He is currently working at the capacity of a Network Architect at Nayna Networks, addressing the data over optical related issues. He has 7 patents pending in the networking area and has many research papers to his credit. He is a member of IEEE, ACM and Phi Kappa Phi. His papers and publications can be found at <http://www.cs.odu.edu/~sudheer>.

Raj Jain: Raj Jain is the Co-founder and Chief Technology Officer of Nayna Networks, Inc. He is currently on a leave of absence from the Ohio State University, Columbus, Ohio, where he is a professor of Computer and Information Sciences. He is a Fellow of IEEE and a Fellow of ACM. He is on the Editorial Boards of [Computer Networks: The International Journal of Computer and Telecommunications Networking](#), [Computer Communications \(UK\)](#), [Journal of High Speed Networks \(USA\)](#), [Mobile Networks and Applications](#), and [International Journal of Wireless and Optical Communications](#) (Singapore). He is currently a Distinguished Lecturer for the IEEE Communications Society and is on Technical Advisory Boards of several companies. He is the author of "Art of Computer Systems Performance Analysis," published by Wiley and winner of the 1991 "Best Advanced How-to Book, Systems" award from Computer Press Association. His second book "FDDI Handbook: High-Speed Networking with Fiber and Other Media" was published in 1994 by Addison Wesley. His papers and publications can be found at <http://www.cis.ohio-state.edu/~jain/>.