Photonic Network Communications, 4:3/4, 219–235, 2002

Invited: Routing Strategies for Capacity-Efficient and Fast-Restorable Mesh Optical Networks

Jean-François Labourdette*, Eric Bouillet, Ramu Ramamurthy, Georgios Ellinas, Sid Chaudhuri, Krishna Bala
Tellium, Inc, 2 Crescent Place, P.O. Box 901, Oceanport, NJ 07757
E-mail: {jlabordette, ebouillet, ramu, gellinas, sc, kbala}@tellium.com

Received October 23, 2001

Abstract. Wavelength division multiplexed (WDM)-based mesh network infrastructures that route optical connections using intelligent optical cross-connects (OXCs) are emerging as the technology of choice to implement the next generation core optical networks. In these architectures a single OXC is capable of switching tens of terabits of traffic per second. With such data transfer rates at stake, it becomes increasingly challenging for carriers to (1) efficiently and cost-effectively operate and manage their infrastructure, and (2) cope with network failures while guaranteeing prescribed service level agreements (SLAs) to their customers. Proper routing of primary and backup paths is a critical component of the routing and restoration architecture required to meeting these challenges.

In this paper we review some of the various strategies and approaches proposed so far to intelligently route connections while at the same time providing guaranteed protection against various types of network failures. We explore the tradeoffs associated with these approaches, and investigate in particular different, sometimes competing aspects, such as cost/capacity required, level of protection (link vs. node failure), restoration time, and complexity of route computation.

Keywords: optical cross-connect (OXC), optical mesh network, routing, restoration

1 Introduction

Dense wavelength division multiplexed (DWDM) mesh networks that route optical connections using optical cross-connects (OXCs) have been proposed as the means to implement the next generation optical networks [1]. Following a wave of timely technological breakthroughs, optical network equipment vendors are now announcing a variety of optical switching systems capable of exchanging and redirecting several terabits of information per second. The dimensions of the proposed switches are colossal, ranging from a few tens to several thousand ports with each single port capable of carrying the equivalent of millions of voice calls, or thousands of video streams. The emergence of new optical technologies is driving down the overall network cost per units of bandwidth, and the trend is accompanied with an explosion of new data service types with various bandwidth characteristics and prescribed quality of service (QoS). Optical network architectures as we envision them now not only provide transmission capacities to higher transport levels, such as inter-router connectivity in an IP-centric infrastructure, but also provide the intelligence required for efficient routing and fast failure restoration in core networks [2–4]. This is possible due to the emergence of optical network elements that have the intelligence required to efficiently manage such networks.

The network architectures under consideration in this paper assume opaque (OEO) switches (with an electronic switch fabric) in an opaque network (with transponders present in the WDM systems). The interfaces to the switch fabric are opaque interfaces, which means that transceivers are present at all interfaces to the switch, and these transceivers provide an OE (input) and EO (output) conversion of the signal. The presence of the transceivers at the edges of the switch fabric enables the switch to access the signal's overhead bytes for control and signaling

* Contact author.

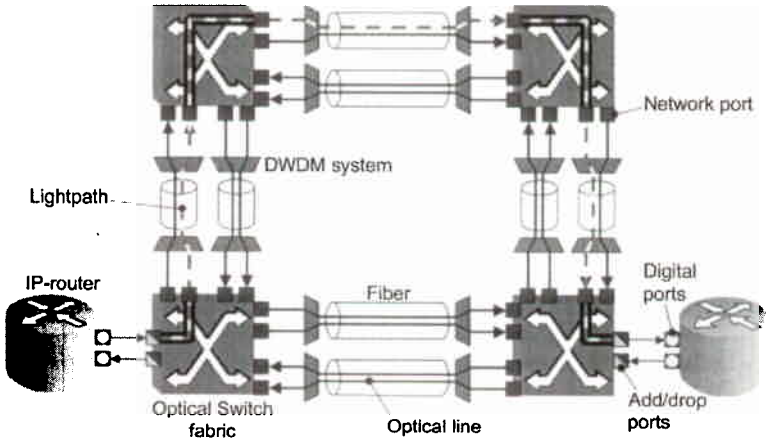

Fig. 1. Optical layer.

functions. The opaque transceivers provide support for fault detection and isolation, performance monitoring, connection verification, neighbor/topology discovery and signaling, as well as support for implementing the network routing and restoration protocols. Furthermore, such a network offers additional key ingredients for a large-scale manageable network: (a) No cascading of physical impairments. This eliminates the need to engineer end-to-end systems and allows full flexibility in signal routing. (b) Multi-vendor interoperability using standard intra-office interfaces. (c) Wavelength conversion enabled. Network capacity can be utilized for service without any restrictions and additional significant cost savings can be offered by sharing restoration capacity in a mesh architecture [5]. (d) The network size and the length of the lightpaths can be large, since regeneration and re-timing is present along the physical path of the signal.

Optical network architectures are exposed to multiple risks of failure, either due to human-induced mishaps such as accidental fiber cuts and operational errors, or equipment malfunctions such as switch and laser failures. Two competing approaches are being proposed for providing the appropriate recovery mechanisms that guarantee service flow continuity in these circumstances. In the peer-to-peer approach [6–8], optical and higher layer equipment operates under the same control plane. In the overlay approach [6,7], optical and logical domains are two separate entities with individual control planes, exchanging management services through a standard interface. The peer-to-peer approach relies on a unified bandwidth management protocol to reassign bandwidth away from defective areas in the network and re-establish the interrupted data services. In the overlay approach, each layer independently relies on its own restoration mechanism in a manner that is independent and transparent to one another. We focus here on the overlay architecture, and specifically on opaque network architectures utilizing OEO switches. Under this architecture, we compare different lightpath on-line routing techniques, as well as different OXC-based protection and restoration approaches. We investigate tradeoffs between those approaches, and discuss reliability vs. complexity and economic considerations.

The optical layer offers and manages the capacity required to transport traffic between clients in the logical layer. Fig. 1 depicts an example of a logical network (two IP routers) linked to an optical network (four optical switches). Optical switch ports are either: (1) add/drop-ports, interfacing the optical layer to the client's logical layer, or (2) network ports, interconnecting optical switches. Using our graph representation, nodes are optical switches, and links are bundles of bi-directional optical channels between pairs of optical switches. An optical channel is a wavelength that connects the network ports of adjacent optical switches. A link in the logical layer is realized by way of optical channels in tandem forming a lightpath (circuit) between the end-nodes of that link. The optical layer faces the same network control and management challenges, and conceptually even borrows solutions from the logical layer (see Rajagopalan et al., Bernstein et al., Hjálmtýsson et al., Sato et al. [6,9–11] for further information and useful references on this topic).

In the overlay approach the layers work individually, and a user network interface (UNI) can specify the exchange of control messages between the two layers. The optical interworking forum (OIF) recently standardized the UNI 1.0 Implementation Agreement [12]. The optical layer can also rely on Generalized MPLS (GMPLS) [13–15] which encompasses all types of architectures, and support wavelength-oriented traffic engineering and management. A carrier will normally acquire network equip-ment from several vendors. The suite of protocols being developed in the Internet engineering task force (IETF) will allow for the seamless interaction between the various network components. As part of the suite, the link management protocol (LMP), for example, is used to maintain control channel connectivity, verify component link connectivity and isolate link, fiber or channel failures within the network [16]. The optical layer can thus rely on neighbor discovery protocol (NDP)/link management protocol (LMP) [7,8,16,17] and open shortest path first (OSPF) [18] to create and publicize the network's topological views. Differences set apart the optical layer from its logical counterpart: (1) routing in the optical layer is exclusively circuit oriented, (2) circuit set-up and tear-down is done at a much slower time scale, and (3) the granularity of the logical layer is much lower than the granularity of the optical layer. The optical layer also supports restoration capabilities for the capacity it manages. The readers can refer to several sources [19–23] for information on failure restoration in a layered architecture. This paper will address restoration in the optical layer exclusively.

The remainder of this paper is organized as follows. Section 3 describes the objectives and constraints that must be considered when routing in the optical layer. Restoration and its relationship to the concept of Shared Risk Group are addressed in Section 4. Section 4 also includes a comparative overview of protection/restoration architectures, as well as specific details on costs, level of protection, speed of restoration, and implementation complexity for a class of dedicated and shared mesh protection/restoration mechanisms. Experiments and results are presented in Section 5, followed by the concluding remarks in Section 6.

2 Routing in the Optical Layer

The procedure to route a lightpath consists of two tasks: (1) route selection, and (2) channel selection. Route selection involves computation of the primary path, and backup path if protection is desired, from the ingress port to the egress port across the mesh optical network. Channel selection deals with selecting individual optical channels on each link along the primary and backup routes. The problems of selecting a route together with selecting channels on the route are closely coupled and if an optimal solution is sought both problems should be solved simultaneously. When solving the route computation problem, several metrics need to be considered. Depending on the allotted budget and the desired QoS, each metric either enters as a parameter in the algorithm's objective function to be minimized, or is used as constraint to eliminate solutions that do not meet practical limits. Some of these metrics are:

1. Cost: the use of optical channels entails a cost. It is henceforth important to ensure that the cumulated cost along a route does not exceed the client's budget.

2. Bit-rate: each optical channel is set to a predetermined bit-rate (e.g., OC-48 or OC-192). The bit-rate of all the selected optical channels along the route must meet the prescribed lightpath bandwidth.

3. Propagation delay: similarly to the cost, delay accumulates along the route. This metric requires a link length or link delay attribute.

The objective of the routing algorithm can be to preserve spare capacity by using minimum number of new optical channels, or to find the solution that incurs the minimum cost. Therefore, among the link attributes that are necessary to carry out routing operations, are (1) the cost per optical channel, (2) the list of available optical channels in the link, (3) their bit-rates, (4) the grooming capabilities of the link,[1] and (5) the propagation delay.

Other factors that must be considered are:

4. Resilience to failure (e.g., link failure vs. node failure, single failure vs. multiple failures): the client expects certain guarantees on the robustness of the connection. However the optical carrier can reduce but not eliminate every risk, such as fiber cuts. It is therefore important to

provision backup capacity on alternate routes where services can be restored if failures occur on the primary path.

5. Restoration time: the service robustness is based on the level of protection offered against a range of failures, but also on the maximum restoration time required to restore the service once a failure occurs. This time depends on the protection architecture, on the characteristics of the path selected by the routing algorithm (e.g., mile-length of the path), and on the network load. These are determining factors that can be considered by the route computation algorithm.

6. Available information about the state of the network. The view of the network varies depending on where the routes are computed. It can be global with the maximum knowledge about network-state and link attributes if the computation is centralized or local with very sparse knowledge if it is distributed across the network. In the latter it may be necessary to produce an educated guess of the route with whatever information is available and defer feasibility and optimality questions—whether all requirements are met—to the moment the lightpath is effectively established through the optical switches.

In this paper, we assume that routing computation is done with access to the complete network information. See Chaudhuri et al. and Bouillet et al. [24,25] for comparison of routing efficiency when only partial information is available.

3 Restoration in the Optical Layer

3.1 Concept of Shared Risk Groups

Failures of multiple optical channels are usually due to fiber or cable cuts. Consider the 6-node optical network of Fig. 2. Each cylinder in the figure represents a conduit. Optical channels across the two links connecting two distinct pairs of nodes traverse the same conduit. If the conduit and the fibers it contains are accidentally severed all the optical channels inside the conduit fail. The concept of shared risk group (SRG) expresses the risk relationship that associates all the optical channels with a single failure

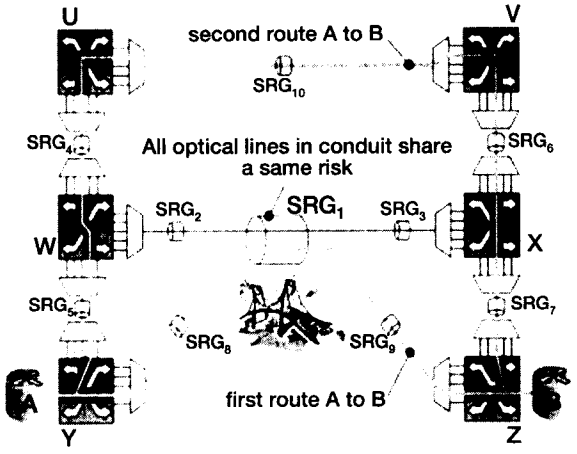

Fig. 2. Shared risk groups.

[23,26,27]. An SRG may consist of all the optical channels in a single fiber, of the optical channels through all the fibers wrapped in the same cable, or of all the optical channels traversing the same conduit. Since a fiber may run through several conduits, an optical channel may belong to several SRGs. Routing algorithms exploit SRG maps to discover SRG-diverse routes so that after any conduit is cut, there is always at least one viable route remaining for restoration. For instance, in Fig. 2 the subtraction of any SRG and the optical channels that traverse it affects at most one of the two routes shown from A to B. For obvious reasons, a network topological view

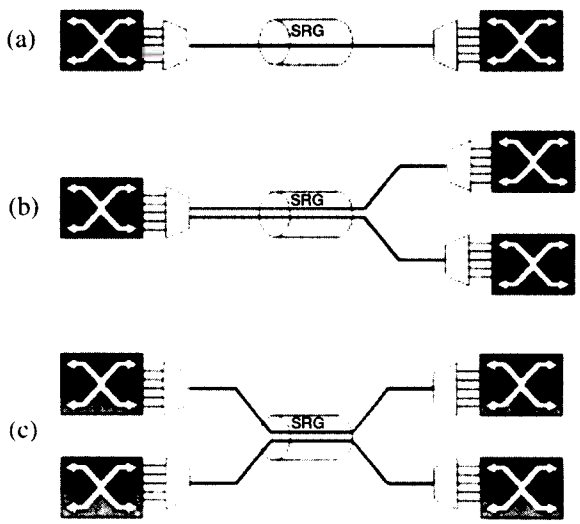

Fig. 3. SRG classification.

alone does not encompass the notion of SRGs. With the exception of the default case, there exists no simple way to automatically generate this information. The network operator must provide it. All SRGs can be expressed as one or a combination of three possible primary types. We describe them in Fig. 3. The default and most conventional type, type (a) in the figure, associates an optical channel risk failure with a fiber cut. Another type of SRG very likely to be encountered, is type (b). This type is typical of fibers terminating at a switch and sharing a same conduit into the office; a conduit cut would affect all the optical channels terminating at the switch. Using elementary transformations for types (a) and (b) [23,24], it is possible to model the network as a graph onto which established routing operations such as shortest-path searches can be applied. Type (c) SRG is the most difficult kind to model and to provide diverse routing for. It occurs in a few instances, such as for example fibers from many origins and destinations routed into a single submarine conduit, or dense metropolitan areas. Contrary to type (a) and (b), there is no convenient way to graphically represent type (c) SRG and their presence can increase dramatically the complexity of the SRG-diverse routing problems [23,26]. By default in this paper all the optical channels between one node-pair belong exclusively to one distinct SRG.

3.2 Node vs. Link Failure

We consider two types of protection: (1) single SRG failure resilience, and (2) single node-and-SRG failure resilience. For protection against several simultaneous failures, see Lumetta and Medard [28]. The problem of insuring survivability against a single failure among a set of lightpaths is addressed in Modiano and Narula-Tam [29]. Resilience against single SRG failure is achieved by way of path diversity, as explained in Section 4.1 and illustrated in Fig. 4. Some level of node failure protection can be realized by way of a redundant switch fabric, but that does not protect from severe events, such as electrical fires that affect both switching fabrics. If protection against this type of failure is also desired, it is necessary to provision routes that are node-and-SRG disjoint, as shown in the example of Fig. 5. However, node-diverse paths consume more resource than the less conservative SRG-diverse scheme pictured in Fig. 4.

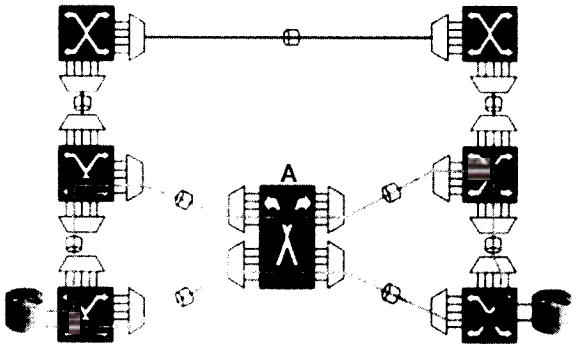

Fig. 4. SRG-diverse paths.

3.3 Taxonomy of Protection/Restoration Architectures

Table 1 presents a classification of protection/restoration architectures in six possible categories (this is an extension of a similar table presented in Doverspike and Yates [30]).

The table enumerates the three components managed during restoration. The components are the restoration route around a failure, the channels used along that route, and the embedding of the route into the switch fabrics. Each category indicates the dependence of each component on the origin of the failure. Components that do not depend on the failure may be assigned before the failure occurs (Categories 1 and 2). For components that are assigned after the failure occurs, the table distinguishes between scenarios with pre-computed routes but without pre-assigned channels[3] (Categories 3 [31,32] and 5), scenarios with pre-computed routes and pre-assigned channels (Category 4), and scenarios where components are determined and assigned after the failure (Category 6) [33]. Categories 4 and 5 depend on the

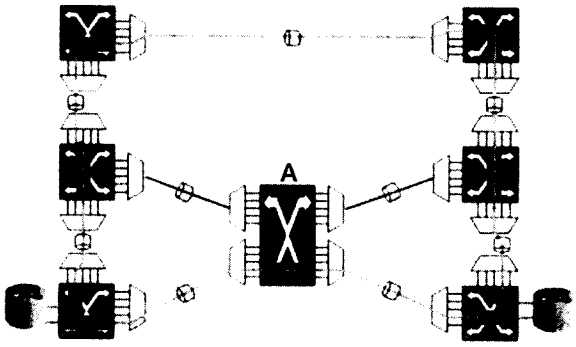

Fig. 5. Node-and-SRG-diverse paths.

Table 1. Different protection/restoration categories and their dependence on failure origins.

Category	Restoration Route		Channel Assignment on Restoration Route		Cross-Connect on Restoration Route	Failure Dependent
	Computed	Assigned	Computed	Assigned		
Dedicated mesh (1 + 1) protection (Cat. 1)	Before	Before	Before	Before	Before	No
Shared mesh restoration with pre-assigned channels (Cat. 2)	Before	Before	Before	Before	After	No
Shared mesh restoration with reserved, not pre-assigned channels (Cat. 3)	Before	Before	N/A	After	After	Yes[2]
Shared mesh restoration with preplanned maps (routes and channels) (Cat. 4)	Before	After	Before	After	After	Yes
Shared mesh restoration with preplanned maps (routes only) (Cat. 5)	Before	After	N/A	After	After	Yes
Reprovisioning (Cat. 6)	After	After	After	After	After	Yes

ability of the optical network to perform rapid fault isolation and select the pre-computed components from a look-up table or map, based on the location of the fault. In the remainder of the paper, we only consider the case of pre-computed restoration paths and pre-assigned channels, where the restoration path is the same independently of the failure (Categories 1 and 2). For the case where the restoration paths are computed in real-time after a failure is detected and localized, see Chao et al. and Grover [34,35]. Protections against failures can also be classified into two main categories, (1) local span protection, and (2) end-to-end or path protection.

3.4 Local Span Protection

In local span protection [36,37], whenever a failure is detected, the optical nodes closest to the failure attempt to reroute the lightpaths through alternate circuits around the failure (Fig. 6). This protection scheme theoretically yields faster and higher availability, since most of the time only the disabled portion of the path is bypassed. On the downside the alternate routes differ for each failure and are difficult to anticipate. The theoretically preferred approach to address this problem is to simulate all possible failures and create maps stored in the optical switches that assign a pre-computed switch configuration for each failure scenario (Category 4 or 5 in Table 1). However, the generated maps may become very large and the concept does not scale well with the size of the network. The approach also entails lengthy computations whenever a new lightpath is provisioned since it must account for every failure scenario

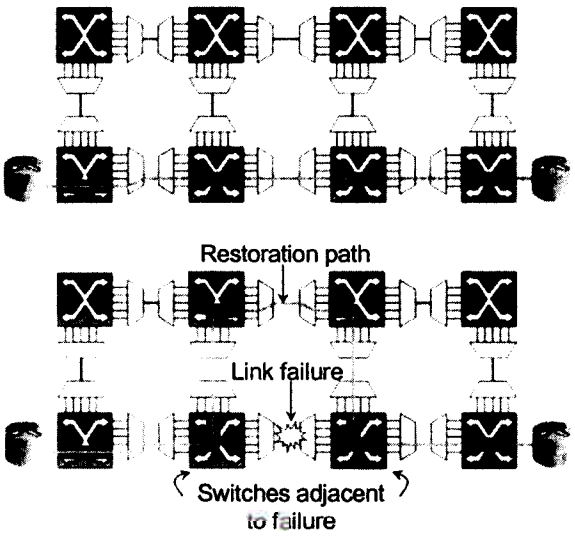

Fig. 6. Local span protection.

in order to populate the map (Categories 4 and 5 in Table 1). For this reason, spare capacity is usually not reserved ahead of time. Instead the restoration routes are computed on the fly upon failure (Category 6 in Table 1). This becomes an issue if the failure disrupts many parallel optical links and sets off a cascade of real-time recovery procedures at the optical switches adjacent to the failure. In either case, this protection scheme relies on the ability of the network to isolate the failure. Finally, link-based restoration schemes require more restoration capacity than path-based schemes [38,39].

3.5 Path Protection/Restoration

In end-to-end or path protection and restoration, the ingress and egress nodes of the failed optical connection attempt to restore the signal on a predefined backup path, which is SRG-disjoint, or diverse, from the primary path [40,41]. Path diversity guarantees that primary and backup lightpaths will not simultaneously succumb to a single failure. Unlike local span protection, secondary routes can be provisioned with the primary routes and thus the restoration may not involve further real-time path computations (Categories 1 to 5 in Table 1). Another advantage of path protection is that the restoration processing can be distributed among ingress and egress nodes of all failed the lightpaths, compared to local span protection where a comparable amount of processing is executed by a smaller set of nodes adjacent to the failure. In the following we will only consider the cases where the protection path is failure-independent and is thus the same for all types of failures. By way of this restriction, the restoration paths may be computed and assigned before failure occurrence (Categories 1, 2, and 3 in Table 1). There are two subtypes of path protection: (1) dedicated mesh (or $1 + 1$) protection and (2) shared mesh restoration.

3.5.1 Dedicated Mesh Protection

(a) Protection against single SRG failures. Dedicated mesh (or $1 + 1$) protection is illustrated in Fig. 7. The network consists of four logical nodes (A to D) and two demands (AB and CD) accommodated across an eight node optical network

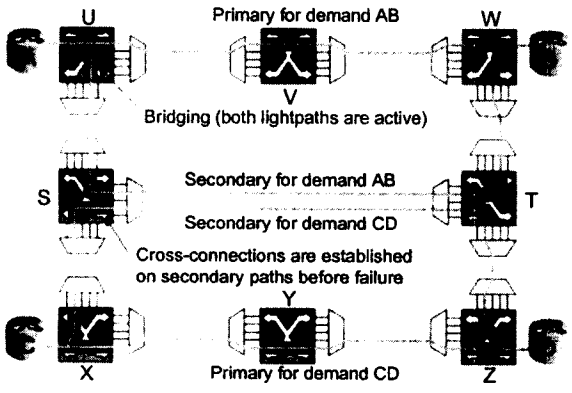

Fig. 7. Dedicated mesh $(1 + 1)$ protection.

(S to Z.) The provisioning algorithm of this architecture computes and establishes simultaneously the primaries and their SRG-disjoint protection paths. During normal operation mode, both paths carry the optical signal and the egress selects the best copy out of the two. In the example of Fig. 7, all the optical channels on primary and secondary paths are active. In particular, the configuration reserves two optical channels between nodes S and T for protection. This is the fastest restoration scheme since for every lightpath one device is responsible for all the necessary failure detection and restoration functions. But it is also the most exigent in terms of resource consumption. The problem of finding SRG diverse routes is trivial if SRGs are of type (a) or (b), or a combination of both, since they can be easily represented in a graph model as indicated earlier in Section 4.1. There exist optimum algorithms to solve this [42,43]. For the case of type (c) SRGs the problem is provable NP-complete [23] and pseudo-optimal solutions must be obtained, using enumerative approaches. The problem is also NP-complete if the selected routes must respect a set of independent constraints, such as maximum round trip-delay (or alternatively maximum path length expressed in geographical distance units) [23].

(b) Protection against single node-and-SRG failures. If protection against node failure is also desired, then primary and backup paths must be node-disjoint in addition to SRG-disjoint. As explained earlier, node protection requires more bandwidth. However, experiments indicate that the two types of dedicated mesh $(1 + 1)$ protection use comparable amount of capacity as will be shown in Section 5.1. The problem of finding node diverse routes is equivalent to the problem of SRG-diverse path routing. The same algorithm is capable of solving either problem, using a graph transformation technique [44] described in Bouillet et al. [23].

3.5.2 Shared Mesh Restoration

(a) Protection against single SRG failures. As in dedicated mesh protection, shared mesh restoration back-up paths are predefined, except that the cross-connections along the paths are not created until a failure occurs. During normal operation modes the spare optical channels reserved for protection are not

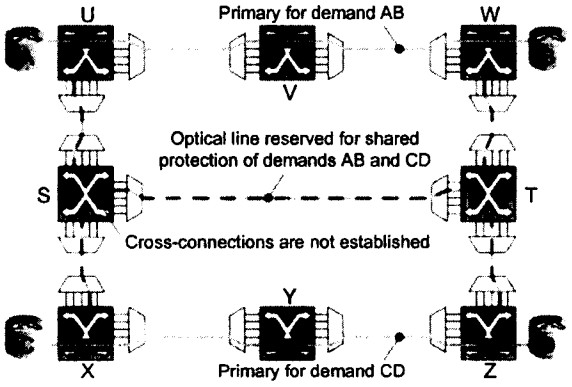

Fig. 8. Shared mesh restoration: before failure.

used. We refer to such channels as reserved (for restoration) shared channels. Since the capacity is only "soft reserved", the same optical channel can be shared to protect multiple lightpaths. There is a condition though that two backup lightpaths may share a reserved channel only if their respective primaries are SRG-disjoint, so that a failure does not interrupt both primary paths. If that happened, there would be contention for the reserved channel and only one of the two lightpaths would be successfully restored. Two lightpaths, or their protection, are said to be mutually compatible, if they are not affected by the same failure. If not, they are incompatible. Fig. 8 (for normal mode) and Fig. 9 (for restoration mode) picture an example of shared mesh restoration. The demand and the network are the same as in dedicated mesh protection. The dashed lines represent reserved channels. Using the routing of Fig. 8, demands AB and CD are compatible with respect to SRG-failures

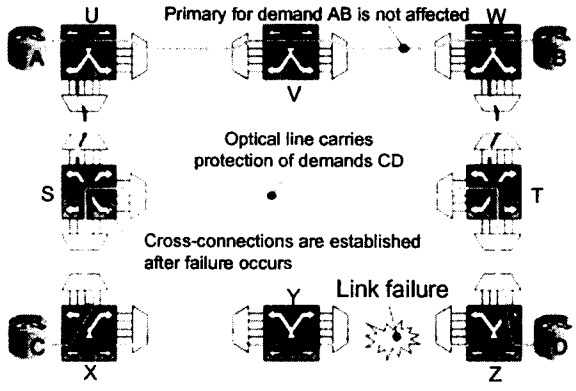

Fig. 9. Shared mesh restoration: after failure.

and thus their protection share a single optical channel in link S-T, one less than in dedicated protection. Upon failure as depicted in Fig. 9, the egress and egress nodes of the disconnected paths (X and Z in example) emit a request to the switches along the protection paths (S and T in example) to establish the cross-connections for that path. Once the cross-connections are established, each ingress and egress node restores the connection to the new path. This architecture requires fewer resources than in dedicated protection. However, the restoration involves a bit more processing to signal and establish the cross-connections along the restoration path and restoration times will be higher than with dedicated mesh protection. There is thus an evident trade-off between capacity utilization (better in the case of shared mesh restoration) and recovery time (better in the case of dedicated mesh protection).

(b) Protection against single node-and-SRG failures. In mesh restoration, node-diversity between primary and backup paths does not guarantee full protection against node failures. Additional sharing restrictions [23] are required to guarantee restoration in case of node failure (for those lightpaths that did not terminate or originate at the failed node).

3.6 Unprotected, Non-Preemptible and Preemptible Lightpaths

If no protection is required for certain types of services, such as best effort services, the connection could simply be routed on a single—unprotected—path. Alternatively, an even more economical solution is to use the reserved channels, which are otherwise idle in normal operation mode, for preemptible service. If necessary, the services routed on this type of connection are interrupted and the reserved channels are freed in order to restore protected lightpaths.

3.7 Characterization of Dedicated and Shared Mesh Protection/Restoration

We characterize in the following table the protection/restoration services presented earlier in this section. For each protection/restoration service and SRG type, we indicate the complexity of the lightpath provisioning operation (Polynomial or NP-complete [45].)[4] We also indicate the speed of protection as well as the cost of the service expressed in amount

Table 2. Summary of different protection/restoration services and their complexity (centralized routing).

Protection/Restoration Services	SRG Type (Section 3.1, Fig. 3)	Complexity of Routing	Restoration Resource (% of Working Capacity)[5]	Speed of Protection
Unprotected preemptible	(a), (b), (c)	Polynomial	Does not apply	Does not apply
Unprotected non-preemptible	(a), (b), (c)	Polynomial	Does not apply	Does not apply
SRG failure dedicated mesh protected	(a), (b) (c)	Polynomial NP-complete	100–170% > 100%	Very fast Very fast
Node-and-SRG failure dedicated mesh protected	(a), (b) (c)	Polynomial NP-complete	100–175% >100%	Very fast Very fast
SRG failure shared mesh restorable	(a), (b), (c)	NP-complete	40–70%	Fast
Node-and-SRG failure shared mesh restorable	(a), (b), (c)	NP-complete	40–80%	Fast

of resources used by the protection mechanism (in percentage of working capacity). Very fast protection/restoration times refer to a few 10s of msec and fast protection/restoration times refer to a few 100s of msec.

4 Experiments and Results

4.1 Resource Utilization for Different Restoration Architectures

All simulation experiments were run on two networks. N17 is a 17-node, 24-edge network that has a degree distribution of (8,6,1,2) nodes with respective degrees (2,3,4,5). N100 is a 100-node, 137-edge network that has a degree distribution of (50,28,20,2) nodes with respective degrees (2,3,4,5). These are realistic topologies representative of existing networks. It is assumed that these networks have infinite link capacity, and that SRGs comprise exclusively all the optical channels in individual links. Five protection/restoration architectures are considered: no protection; dedicated mesh protection with guaranteed recovery from single link failures; dedicated mesh protection with guaranteed recovery from single link or node failures; shared mesh restoration with guaranteed recovery from single link failures; shared mesh restoration with guaranteed recovery from single link or node failures. In N17, demand is uniform, and consists of two bi-directional lightpaths

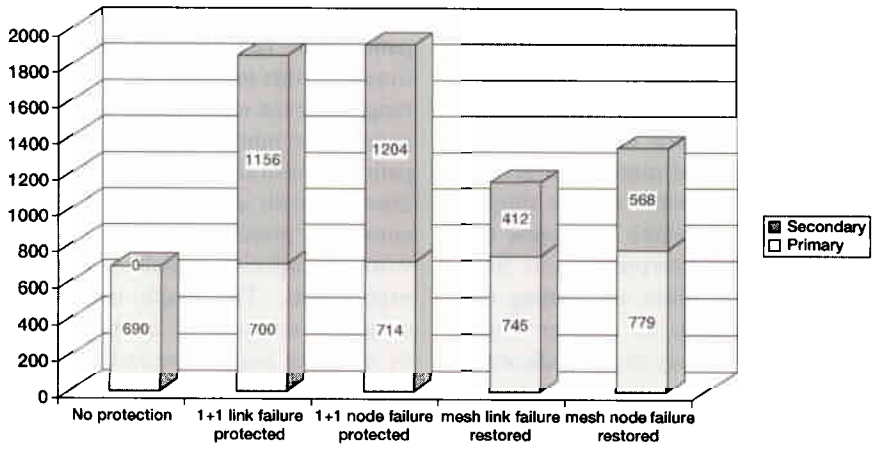

Fig. 10. Comparison of capacity usage for different protection/restoration architectures (17 nodes).

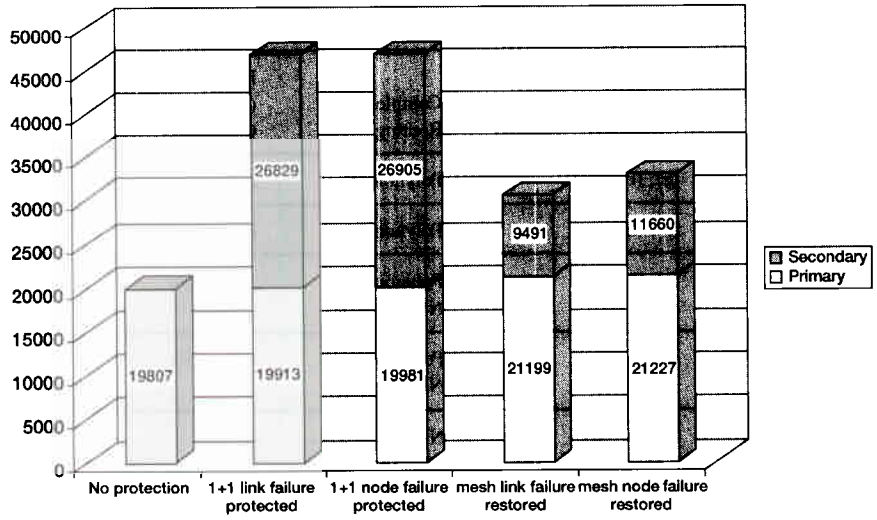

Fig. 11. Comparison of capacity usage for different protection/restoration architectures (100 nodes).

between every pair of nodes. That amounts to 272 lightpaths. In N100, 3278 node-pairs out of 4950 possible node pairs are connected by one bi-directional lightpath. Requests for lightpaths arrive one at the time (on-line routing) in a finite sequence and in an order that are arbitrary but common to each scenario to ensure a fair comparison. Figures of merit are capacity requirements separated into their primary and protection/restoration parts, and expressed in units of bi-directional OC-48 channels. Results are presented in Figs. 10 and 11. The quantities shown on the charts are averaged over series of 10 experiments using various demand arrival orders. These results indicate that link-disjoint and node-disjoint dedicated mesh protection approaches[6] consume approximately the same total amount of capacity. This is expected since dedicated protection against link failures protects against node failures as well for nodes up to and including degree 3 and these nodes constitutes a majority of the nodes in these two networks. This property does not apply to shared mesh restoration, and lightpaths that must be protected against single node failures, even for nodes of degree 3. Because of this and other reasons whose interpretation is not within the scope of this presentation, the routing of shared mesh restorable lightpaths to recover from single node failures takes (relatively to the dedicated schemes) more resources than to recover from single link failures. The difference however should be considered in light of the benefits of protecting the network against single node failures.

4.2 Trading-off Backup Path Length vs. Protection Capacity

While longer back-up paths allow re-using existing shared channels and therefore minimize the total capacity required, they have the potential to increase the restoration time as restoration signaling traverses more nodes and links. In this section, we explore trade-offs between back-up path lengths and total capacity required to route the demand.

4.2.1 Limiting the Length of the Back-up Path

In shared mesh restoration, provisioning of protection paths sometimes requires longer paths that consist exclusively of shared channels, rather than shorter paths where new channels must be reserved. Fig. 12 illustrates this tendency on a 17-node network over a range of seven experiments numbered from 0 to 8. In experiment number j the length of each protection path is constrained to be at most the shortest-hop alternate path plus j hops. The plots indicate the amount of protection capacity as a percentage of the working capacity, which does not vary across the experiments. The graph indicates that substantial capacity savings, 17% in this example, are possible by allowing longer restoration paths. The graph also indicates that four hops over shortest-hop alternate paths are sufficient to gain most of the benefits. In these experiments, the average restoration hop-distance increases 25% from four to five hops.

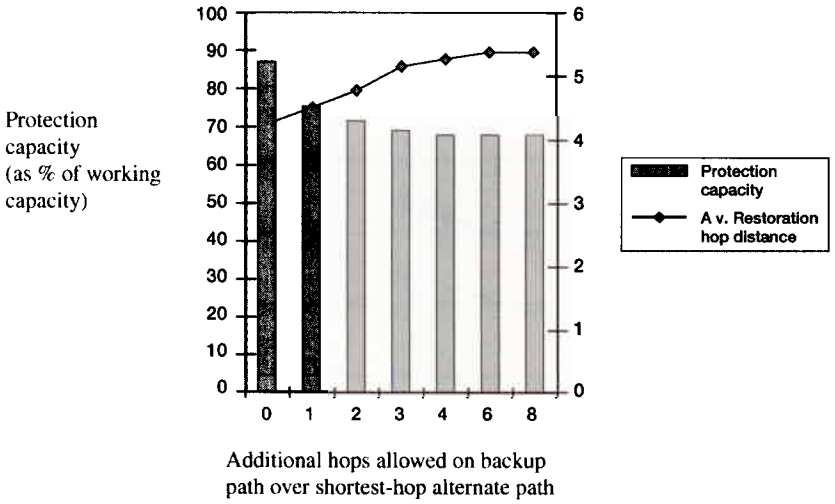

Fig. 12. Effect of additional hops allowed on backup path over shortest-hop alternate path.

4.2.2 Changing the Cost of Shared Channels

The metric or policy used for weighting the edges should be different for primary paths and backup paths. For primary paths it is the real cost w_e of using the edges. For backup path it is a function of the primary path where an edge e is assigned (1) infinite weight if it intersects with an SRG of the primary path, (2) weight w_e if new capacity is required to provision the path, and (3) weight $s_e \leq w_e$ if the path can share existing capacity reserved for pre-established backup paths. Quite evidently, the underlying idea here is to encourage "sharing", whereby existing capacity can be reused for provisioning multiple backup paths. The condition for sharing is that the backup paths must not be activated simultaneously, or in other words that their respective primaries must be pair-wise SRG-disjoint so that they do not fail simultaneously. The ratio s_e to w_e, $\varepsilon = s_e/w_e$ can be adjusted for the desired level of sharing. For smaller values of s_e, backup paths will be selected with the minimization of the number of non-shareable edges (weights w_e) in view, eventually leading to arbitrary long paths (as expressed in number of hops) that consist uniquely of shareable edges (weights s_e.) For larger values of s_e routing is performed regardless of sharing opportunities and backup paths will end-up requiring substantially more capacity (see also Doucette et al. and Bouillet et al. [46,47]). Another routing variation is based on the realization that, everything else being equal, it should be better to use a new channel as part of the back-up path (where it can potentially be shared in the future) rather than as part of the primary path (where it cannot). One can assign a weight t_e, $s_e \leq t_e \leq w_e$, close to w_e, to an edge e of the back-up path if new capacity is required. The ratio of t_e to w_e can be adjusted, in conjunction with the ratio of s_e to w_e, to achieve desired results such as reduced capacity requirements [48]. In this section, we study various values of s_e with respect to w_e. The algorithm is exercised on randomly generated net-

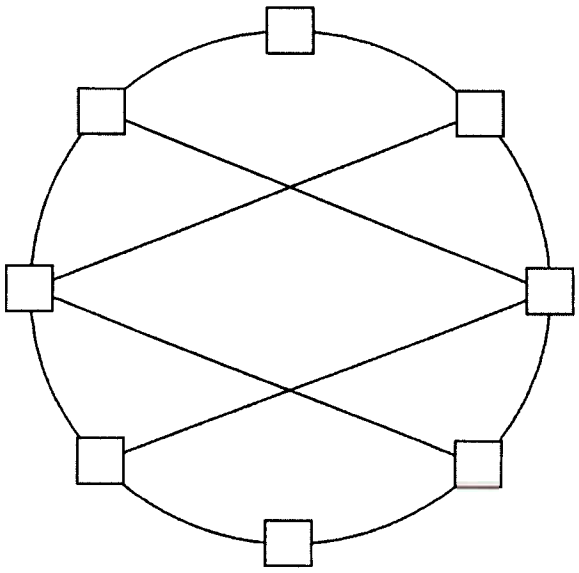

Fig. 13. Example of chordal-ring network.

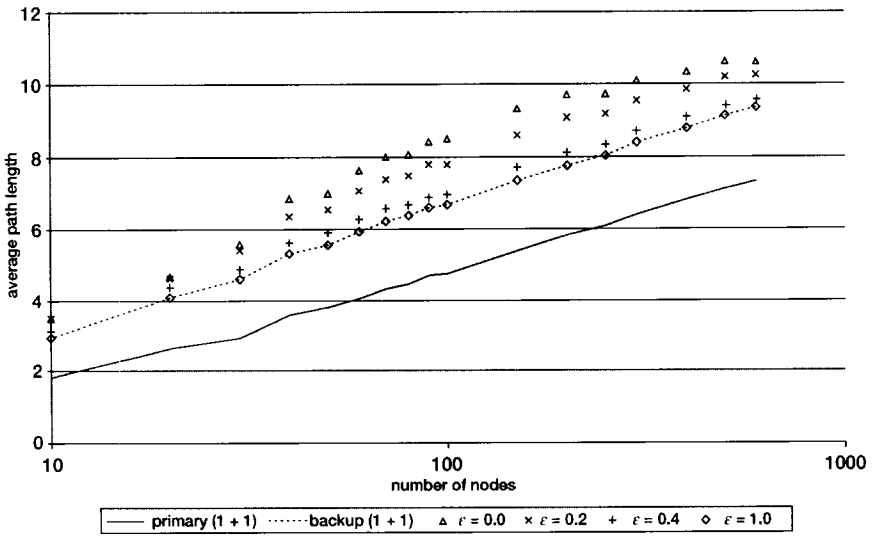

Fig. 14. Effects of shareable channel cost on average path length, degree 3 random chordal-ring networks.

works and real-life networks varying in size and degree.[7,8]

Experiments—randomly generated chordal-ring networks. In this set of experiments we investigate mesh restoration in chordal-ring networks loaded with full demand connectivity. Chordal-ring networks consist of nodes arranged and connected in a ring fashion, and further connected by "chord"-edges across the ring (see Fig. 13).

We chose chordal-rings for their protection availability, i.e., for every path it is possible to find a corresponding backup path that is edge (and in fact node) disjoint. For this study we implemented a tool to randomly generate chordal-rings, with the ability to specify number of nodes and degree distributions (edges per node.) In this set of experiment we assess the effect of $\varepsilon = s_e/w_e$, the shareable channel cost to non-shareable channel cost ratio, on randomly generated chordal-ring networks.

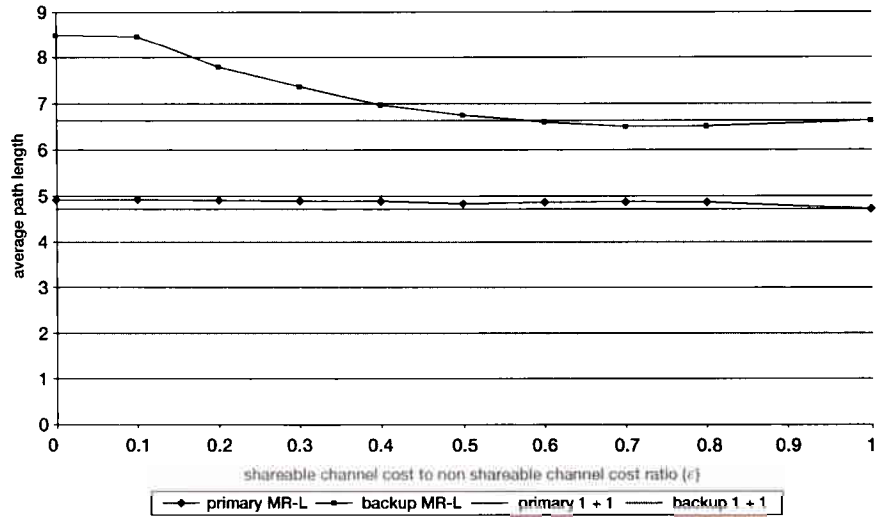

Fig. 15. Effects of shareable channel cost on average path length, 100 nodes, and 150 edges random network.

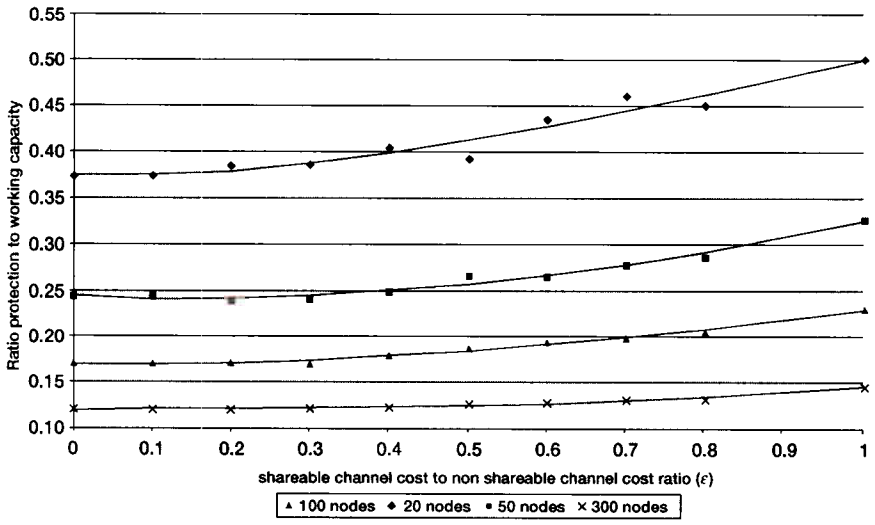

Fig. 16. Effect of shareable channel cost on protection to working capacity ratio, for 20 to 300 nodes degree 3 random networks.

Fig. 14 shows the average path length of the primary and back-up paths of dedicated mesh (1 + 1) protected lightpaths, and the average path length of the back-up path of shared mesh restorable lightpaths with different values of ε. As expected, as ε gets closer to one, the average back-up path length for shared mesh restoration becomes equal to that of dedicated mesh protection.

Fig. 15 shows the average path length of primary and back-up paths of shared mesh restorable (MR-L) lightpaths as ε varies from zero to one, compared to the corresponding average path lengths for dedicated mesh (1 + 1) lightpaths.

Fig. 16 shows the evolution of the ratio of protection to working capacity for different size networks as ε varies from zero to one. While the average length of back-up paths for shared mesh restoration with $\varepsilon = 1$ and dedicated mesh protection is the same as shown in Figs 14 and 15, the total capacity required is not the same. Sharing of channels can be accomplished on the back-up path with shared mesh restoration even if the computation of that path

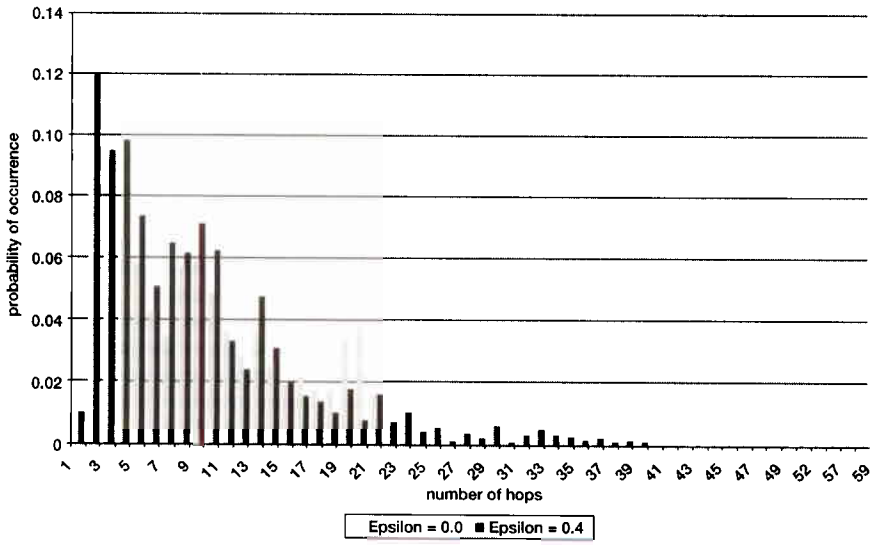

Fig. 17. Effect of shareable channel cost on back-up path length distribution.

is blind to any potential sharing as is effectively the case when $\varepsilon = 1$.

Experiment—200+ node, 300+ edge network. In Fig. 17, we show the distribution of backup path lengths for values of $\varepsilon = 0.0$ and $\varepsilon = 0.4$ for a network with 200+ nodes and 300+ edges. The corresponding average backup lengths for these values of ε were respectively 13.26 hops and 8.875 hops. Value $\varepsilon = 0.4$ incurs a 2.5% increase of total capacity in comparison to $\varepsilon = 0.0$.

In summary, we observe a significant improvement in terms of number of hops for the backup paths when $s_e \geq 0.1w_e$. In the range 0 to $0.5w_e$ the capacity utilization remains within the order of a few percentages of its minimum. The improvement in terms of number of hops becomes marginal for $s_e \geq 0.5w_e$, whereas increase of capacity utilization becomes noticeable. Based on these experimental findings, it appears that s_e in the range $0.1w_e$ to $0.5w_e$ yields a good compromise for different tested networks, random and real-life.

5 Conclusion

In this paper we have reviewed a set of protection/restoration architectures as well as various techniques for routing lightpaths in mesh optical networks. In order to conduct these comparisons we first described multiple scenarios based on architecture characteristics (e.g., shared risk groups), and the type of failures for which protection was required (node or link). We then compared several specific protection/restoration mechanisms taking into consideration figures of merits such as capacity requirements, robustness to different failure modes, and speed of restoration.

It is well known that dedicated mesh $(1 + 1)$ protection stands-out in terms of restoration speed, and protection availability for all types of single failures and network architectures. On the other hand, our experimental results show that this protection mechanism consumes substantially more resources to provide the required level of protection than other types of restoration, in particular than shared mesh restoration. If minimizing the cost of the service is the main objective, then shared mesh restoration, which allows sharing of reservation-channel among the restoration-routes, is more appropriate. Besides capacity savings, shared mesh restoration has the additional advantage over dedicated mesh protection that it can support pre-emptible services that use idle restoration channels. We also observed that the amount of resources used with dedicated mesh protection when protecting against single node failure is only marginally higher than the amount used when protecting against single node failure is only marginally higher than the amount used when protecting against single SRG failure. Unlike dedicated mesh protection, shared mesh restoration protecting against single node failures consumes substantially more resources than shared mesh restoration protecting against single link failures.

Other important aspects of routing and restoration are constraints on end-to-end delays and on restoration times. With shared mesh restoration in particular, but restoration in general, the restoration time increases with the length of the back-up or restoration path. It may be therefore important to limit the length of the restoration paths in order to (1) respect prescribed end-to-end delays and provide equivalent level of service on primary and back-up paths, as well as (2) insure fast restoration times within defined SLAs. The paper also presents a framework and experimental results for trading off back-up path length against restoration capacity. It is shown that this trade-off can be controlled through certain parameters to achieve a desired set of results.

Notes

1. That is, the link ability to multiplex sub-rate lightpaths into higher rate channels.
2. The dependency on the failure is only for the channel assignment, not the route determination.
3. The channels are not pre-computed (the routes are) but are assigned during the restoration event from a shared pool. The routing is such that there are enough channels available to restore paths for any single failure.
4. NP-completeness proofs can be found in Bouillet et al. [23].
5. Note that all the percentages mentioned in this table are based on a particular set of experiments.
6. Which insures guaranteed recovery against single link, respectively node and link, failures.
7. In all networks we assumed that all links have unit cost $w_e = 1\ \forall e$ (i.e., not distance sensitive).
8. We did not account for capacity constraints in our experiments and assumed unlimited capacity. In a real network with capacity constraints, we expect the value of s_e to have a lesser impact on backup path lengths.

References

[1] T. E. Stern, K. Bala, Multiwavelength optical networks: A layered approach (Prentice Hall, May 1999).

[2] D. Benjamin, R. Trudel, S. Shew, E. Kus, Optical services over an intelligent optical network, IEEE Communications Magazine, vol. 39, no. 9, (Sept. 2001), pp. 73–78.

[3] E. Varma, S. Sankaranarayanan, G. Newsome, Z. Lin, H. Epstein, Architecting the services optical network, IEEE Communications Magazine, vol. 39, no. 9, (Sept. 2001), pp. 80–87.

[4] A. McGuire, S. Mirza, D. Freeland, Application of control plane technology to dynamic configuration management, IEEE Communications Magazine, vol. 39, no. 9, (Sept. 2001), pp. 94–99.

[5] S. Chaudhuri, E. Bouillet, G. Ellinas, Addressing Transparency in DWDM Mesh Survivable Networks, Proc. of OFC 2001, paper Tu05, Anaheim, CA (March 2001).

[6] B. Rajagopalan et al., IP over Optical Networks: A Framework, draft-many-ip-optical-framework-03.txt, IETF draft, (Jan. 2001).

[7] B. Rajagopalan, D. Pendarakis, D. Saha, S. Ramamurthy, K. Bala, IP over optical networks: Architectural aspects, IEEE Communications Magazine, vol. 38, no. 9, (Sept. 2000), pp. 94–102.

[8] G. Bernstein, J. Yates, D. Saha, Control and management of optical transport networks, IEEE Communications Magazine, vol. 38, no. 10, (Oct. 2000), pp. 94–102.

[9] G. Hjálmtýsson, J. Yates, S. Chaudhuri, A. Greenberg, Smart routers—simple optics: An architecture for the optical internet, IEEE/OSA Journal of Lightwave Technology, vol. 18, no. 12, (Dec. 2000), pp. 1880–1891.

[10] G. Bernstein, J. Yates, D. Saha: IP-centric control and management of optical transport networks, IEEE Communications Magazine, vol. 38, no. 10, (Oct. 2000), pp. 161–167.

[11] K. Sato, S. Okamoto, Photonic transport technologies to create robust backbone networks, IEEE Communications Magazine, vol. 37, no. 8, (Aug. 1999), pp. 78–87.

[12] B. Rajagopalan et al., User Network Interface (UNI) 1.0 Signaling Specification, OIF Implementation Agreement OIF-UNI-01.0, (Oct. 2001).

[13] E. Mannie et al., Generalized Multi-Protocol Label Switching Architecture, draft-many-gmpls-architecture-01.txt, IETF draft, (Nov. 2001).

[14] L. Berger et al., Generalized MPLS—Signaling Functional Description, draft-ietf-mpls-generalized-signaling-07.txt, IETF draft, (Nov. 2001).

[15] L. Berger et al., Generalized MPLS Signaling—RSVP-TE Extensions, draft-ietf-mpls-generalized-rsvp-te-06.txt, IETF draft, (Nov. 2001).

[16] J. Lang et al., Link Management Protocol (LMP), draft-ietf-ccamp-lmp-02.txt, IETF draft, (Nov. 2001).

[17] A. Fredette et al., Link Management Protocol (LMP) for DWDM Optical Line Systems, draft-fredette-lmp-wdm-03.txt, IETF draft, (Nov. 2001).

[18] J. Moy et al., OSPF version 2, IETF RFC 2328, (April 1998).

[19] P. Demeester et al., Resilience in multilayer networks, IEEE Communications Magazine, vol. 37, no. 8, (Aug. 1999), pp. 70–76.

[20] S. Makam et al., Framework for MPLS Based Recovery, draft-ietf-mpls-recovery-frmwrk-03.txt, IETF draft, (July 2001).

[21] J. Manchester, P. Bonenfant, C. Newton, The evolution of transport network survivability, IEEE Communication Magazine, vol. 37, no. 8, (Aug. 1999), pp. 44–51.

[22] R. Doverspike, S. Phillips, J. Westbrook, Transport network architectures in an IP world, Proc. of IEEE Infocom 2000, Tel Aviv, Israel, vol. 1, (March 2000), pp. 305–314.

[23] G. Ellinas, et al., Routing & Restoration Architectures in Mesh Optical Networks, to appear in Optical Network Magazine.

[24] S. Chaudhuri, R. Ramamurthy, S. Sengupta: Comparison of Centralized and Distributed Provisioning of Lightpaths in Optical Networks, Proc. of OFC 2001, paper MH4, Anaheim, CA, March 2001.

[25] E. Bouillet, J.-F. Labourdette, G. Ellinas, R. Ramamurthy, S. Chaudhuri, Stochastic approaches to route shared mesh restored lightpaths in optical mesh networks, Proc. of IEEE Infocom 2002, New York, NY, June 2002.

[26] J. Strand, A. Chiu, R. Tkach, Issues for routing in the optical layer, IEEE Communications Magazine, vol. 39, no. 2, (Feb. 2001), pp. 81–87.

[27] R. Ramamurthy et al., Capacity performance of dynamic provisioning in optical networks, IEEE/OSA Journal of Lightwave Technology, vol. 19, no. 1, (Jan. 2001), pp. 40–48.

[28] S. Lumetta, M. Medard, Towards a deeper understanding of link restoration algorithms for mesh networks, Proc. of IEEE Infocom 2001, vol. 1, Anchorage, Alaska, (April 2001), pp. 367–375.

[29] E. Modiano, A. Narula-Tam, Survivable routing of logical topologies in WDM networks, Proc. of IEEE Infocom 2001, vol. 1, Anchorage, Alaska, (April 2001), pp. 348–357.

[30] R. Doverspike, J. Yates, Challenges for MPLS in optical network restoration, IEEE Communication Magazine, vol. 39, no. 2, (Feb. 2001), pp. 89–96.

[31] S. Datta, S. Sengupta, S. Biswas, S. Datta, Efficient channel reservation for backup paths in optical mesh networks, Proc. of IEEE Globecom 2001, vol. 4, San Antonio, TX, (Nov. 2001), pp. 2104–2108.

[32] G. Austin et al., Fast, scalable, and distributed restoration in general mesh optical networks, Bell-Labs Technical Journal, vol. 6, no. 1, (Jan.–June 2001).

[33] B. Doshi, S. Dravida, P. Harshavardhana, O. Hauser, Y. Wang, Optical network design and restoration, Bell-Labs Technical Journal, vol. 4, no. 1, (Jan.–March 1999), pp. 58–84.

[34] C.-W. Chao, P. M. Dollard, J. E. Weythman, L. T. Nguyen, H.Eslambolchi: FASTAR—A robust system for fast DS3 restoration, Proc. of IEEE Globecom 1991, (Phoenix, AZ, Dec. 1991), pp. 1396–1400.

[35] W. D. Grover, in Distributed restoration of the transport network, in Telecommunications Network Management into the 21st Century—Techniques, Standards, Technologies and Applications, eds. S. Aidarous, T. Plevyak (IEEE Press, 1994).

[36] B. O. Venables, W. D. Grover, M. H. MacGregor, Two strategies for spare capacity placement in mesh restorable networks, Proc. of IEEE ICC 1993, (Geneva, Switzerland, May 1993), vol. 1, pp. 267–271.

[37] M. Herzberg, S. Bye, An optimal spare-capacity assignment model for survivable networks with hop limits, Proc. of IEEE Globecom 1994, (New orleans, LA, May 1994), pp. 1601–1607.

[38] R. Iraschko, M. H. MacGregor, W. D. Grover, Optimal capacity placement for path restoration in STM or ATM mesh survivable networks, IEEE/ACM Trans. on Networking, vol. 6, no. 3, (June 98), pp. 325–336.

[39] R. R. Iraschko, M. H. MacGregor, W. D. Grover, Optimal capacity placement for path restoration in mesh survivable networks, Proc. of ICC 1996, Dallas, TX, (June 1996), IEEE ICC, vol. 3, pp. 1568–1574.

[40] S. Ramamurthy, B. Mukherjee, Survivable WDM mesh networks. Part I—Protection, Proc. of IEEE Infocom 1999, (New York, NY, March 1999), vol. 2, pp. 744–751.

[41] S. Ramamurthy, B. Mukherjee, Survivable WDM mesh networks, Part II—Restoration, Proc. of IEEE ICC 1999, (Vancouver, Canada, June 1999), vol. 3, pp. 2023–2030.

[42] J. W. Suurballe, Disjoint Paths in a Network, Networks, vol. 4, no. 2, (Apr.–June 1974), pp. 125–145.

[43] R. Bhandari, Survivable Networks: Algorithms for Diverse Routing (Kluwer Academic Publishers, 1999).

[44] G. Chartrand, L. Lesniak, Graphs & Digraphs (Chapman and Hall, 1996).

[45] M. Garey, D. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness (W. H. Freeman and Company, 1979).

[46] J. Doucette, W. Grover, T. Bach, Bi-criteria studies of mesh network restoration: path length vs. capacity tradeoffs, Proc. of OFC 2001, paper TuG2, Anaheim, CA (March 2001).

[47] E. Bouillet, J.-F. Labourdette, R. Ramamurthy, S. Chaudhuri, Enhanced algorithm cost model to control tradeoffs in provisioning shared mesh restored lightpaths, Proc. of OFC 2002, paper ThW2, Anaheim, CA (March 2002), pp. 544–546.

[48] E. Bouillet, J.-F. Labourdette, R. Ramamurthy, Impact of link cost assignment on routing in optical mesh networks, Internal Tellium Report, (2001).

Dr. Jean-Francois Labourdette is manager of network routing and design at Tellium, a position he has held since October 2000. In this position, he is responsible for Tellium routing architecture and algorithms, and network dimensioning and design studies.

Previously, he was a manager of multi-service globalization planning at AT&T, where he was responsible for network, service and operations planning for AT&T international data services (Frame Relay, ATM, and MPLS IP VPN). Earlier at AT&T, he was a system engineer in the routing planning group, working on dynamic call routing for AT&T switched network and facility routing and re-arrangement for AT&T network.

Jean-Francois holds a Ph.D from Columbia University, where he was the recipient of the 1991 Eliahu I. Jury Award for best dissertation. He is a senior member of the IEEE and a member of the Optical Society of America.

Dr. Eric Bouillet is a senior network architect at Tellium. In this role, he works on the design of optical networks and optimization of lightpath provisioning and fault restoration algorithms.

Before joining Tellium, Eric was a member of the technical staff in the Mathematical Sciences Research center in Bell Labs, Lucent Technologies, Murray Hill, New Jersey, working in the areas of routing in ATM and Optical networks.

Eric holds an MS and a Ph.D. in Electrical Engineering from Columbia University. He also holds a joint degree from l'Ecole Nationale Supérieure des Télécommunications ENST Paris and EURECOM Sophia Antipolis. He has authored and co-authored several journal and conference papers and has a number of U.S. and international patents pending in the area of optical networking.

Dr. Ramu Ramamurthy is a senior network architect at Tellium, where he works on the design of algorithms and protocols for dynamic provisioning and restoration in optical networks.

Prior to joining Tellium, he was a research scientist at Telecordia Technologies, where he worked on network control and the management of IP/WDM optical networks.

Ramu holds a B.Tech. degree from the Indian Institute of Technology, Madras, and MS and Ph.D. degrees from the University of California, Davis.

Dr. Georgios Ellinas is a senior network architect at Tellium. In this role, he works on lightpath provisioning and fault restoration algorithms in optical mesh networks, and the architecture design of the company's MEMS-based Aurora Full-Spectrum Optical Switch.

Before joining Tellium, George was a member of the Optical Networking Research Program at Telcordia Technologies (formerly Bellcore), performing research for the Optical Networks Technology Consortium (ONTC) and Multiwavelength Optical Networking (MONET) projects, from 1993 to 2000. In 1999, he became a senior research scientist in Telcordia's Optical Networking Research Group.

George also served as an adjunct assistant professor at Columbia University and the University of Maryland, teaching courses on multiwavelength optical networking in 1999 and 2000, respectively.

George holds a BS, an MS and a Ph.D. in Electrical Engineering from Columbia University. He was awarded a Fulbright fellowship, from 1987 to 1991, for undergraduate studies at Columbia

University and has authored and co-authored more than 35 journal and conference papers. George is also the holder of eight U.S. Patents on optical networking and has more than 25 U.S. and international patent applications currently pending.

Dr. Sid Chaudhuri is director of network architecture at Tellium, a position he has held since March 2000. In this role, he led the development of the mesh restoration and IP-centric control architecture implemented in the company's Aurora Optical Switch. He also led the development of Tellium's StarNet Software Design Tools. Prior to joining Tellium, Sid worked for AT&T Laboratories Research where he developed intelligent optical network architectures using SONET, DWDM and optical switching technologies and spearheaded AT&T's Core Transport Network architecture. Previously, he was a distinguished member of the technical staff at Bell Laboratories, where he developed SONET/SDH transport and cross-connect system architecture.

Sid is President and Chairman of the Optical Internetworking Forum (OIF) and served as chairman of its physical and link layer group, where he led the group in the first set of interoperability agreements. Sid holds a Ph.D. from the University of Pittsburgh, where he was an Andrew W. Mellon fellow.

Dr. Krishna Bala is a co-founder and chief technical officer of Tellium. He is the lead system architect for Tellium's Aurora Optical Switch and the Full Spectrum All-Optical Switch. From 1997 to 1999, Krishna was Tellium's manager of optical cross-connect product development. He was a member of the team that launched Tellium's optical switching product line and established the feasibility of key technologies. Krishna is a pioneer in the field of optical switching and networking, in which he is the holder of several patents. He is also a co-author of the book *Multiwavelength Optical Networks: A Layered Approach* (Prentice Hall, 1999).

Prior to joining Tellium, Krishna was a senior scientist in the Optical Networking Group of Bellcore, a research and telecommunications services provider, from 1992 to 1997. While at Bellcore, he was responsible for local exchange optical networking architecture development and analysis and received the Bellcore President's Award for outstanding achievement. Krishna holds a Ph.D. in Electrical Engineering from Columbia University. His thesis at Columbia was one of the first in the area of optical wavelength routing.

Krishna is chairman of the signaling working group of the Optical Internetworking Forum that is an industry group that enables the creation of optical networking standards.