Use of Backup Trees to Improve Resource Efficiency of MPλS Recovery Mechanisms

A. Groebbens, D. Colle, S. Demaesschalck, C. Develder, M. Pickavet, P. Demeester
Department of Information Technology, The Ghent University - IMEC
Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium.
email: {adelbert.groebbens, dcolle, sophie.demaesschalck, chris.develder, mario.pickavet, demeester}@intec.rug.ac.be

ABSTRACT:

Multi-Protocol Lambda Switching (MPλS), the MPLS "label switching paradigm" applied to the optical networking domain, is a promising solution for the dynamic control of Optical Transport Networks (OTNs). To protect the huge amount of traffic carried by an OTN against network failures, MPLS protection techniques (based on pre-established backup paths) can be applied with some modifications. Unfortunately, in an MPλS network, these protection techniques suffer from the fact that the capacity they consume cannot be shared among different single network failures (leading to so-called dedicated protection). In order to reduce the amount of consumed protection capacity, a novel potential improvement based on "backup trees" is introduced. It will be investigated quantitatively how much capacity savings are to be expected by the use of backup trees.

KEYWORDS:

IP-over-OTN, MPLS, MPλS, Recovery, Capacity Dimensioning, Integer Linear Programming, Backup Tree

Introduction : IP-over-OTN and MPλS

Due to the fast increase of IP traffic and the enormous bandwidth potential of (all-)optical transport networks, an IP-over-OTN network scenario is likely to be widespread in future communication networks.

This evolution is fuelled by two main drivers. Firstly, the current dominant amount of data traffic (which still keeps growing) makes it more-and-more efficient to transport all services over IP. IP will be the common revenue-generating convergence client layer. Secondly, to cope with the need for ever increasing capacity, the electrical layer will have to be bypassed. Indeed, operators have been introducing Wavelength Division Multiplexing (WDM) into their backbone network, upgrading point to point links by using multiple channels (wavelengths) on one fiber. For the thus created (very) high capacity links, electronic switching becomes the bottleneck. As a result, in order to keep the load on the IP routers reasonable, optical bypassing (feasible with Optical CROSS-Connects (OXCs)) becomes a necessity ([ghani], [manzalini]). Implementing network functionality directly in the optical layer by using OXCs, leads to a so-called Optical Transport Network (OTN). This first part of the network evolution is shown in Figure 1: OXCs allow that lightpaths through an OTN can be setup end-to-end. However, these paths still have to be provisioned manually, which often takes a long time. Besides that, more and more flexibility is required by the clients and it becomes a necessity to allow the client layer to initiate this procedure (see also Figure 1). Appropriate protocols are needed for the automatization and speed up of this process. Multi-Protocol Lambda Switching (MPλS) is such a candidate ([ghani], [awduche]). MPλS is an extension of the MPLS paradigm (developed in the Internet Engineering Task Force (IETF)) applied to an optical (transport) network. MPλS controls the OXCs, which then are called Optical Label Switched Routers (O-LSRs, see Figure 2). The main idea is to consider each wavelength on an OTN-link as an MPLS label. The label processing is done optically, by configuring the OXCs in the network correctly. MPλS relies on similar protocols as MPLS does, and has, for example, a built-in protocol to distribute the labels (e.g. Label Distribution Protocol (LDP)), i.e. to negotiate about the OXC configuration for the lightpath to be setup.

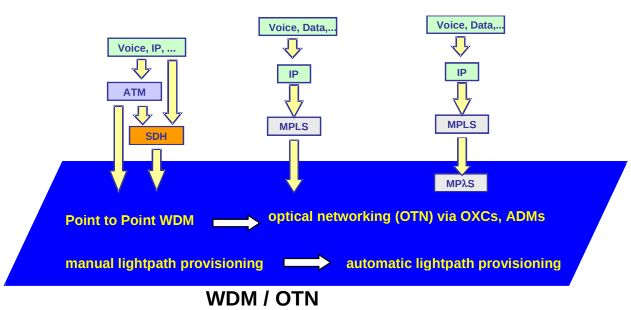


Figure 1: evolution towards IP over OTN scenario

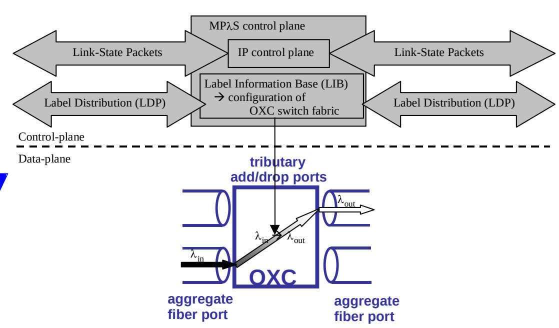


Figure 2: overview of an MPλS capable OXC, showing the data and control planes

At the same time, it is of paramount importance for mission-critical IP-over-OTN networks to be able to recover quickly from frequently occurring network failures. Network recovery in MPLS and MPλS will be discussed in the next section.

MPLS and MPλS protection

Certainly, there is a need for recovery in the IP layer. Optical layer recovery schemes cannot recover for example from a failing IP router. Furthermore, a node failure in the OTN may isolate a connected IP router and the recovery of IP traffic transiting the now isolated router can only be realized by rerouting in the IP layer. The OSPF routing protocol has inherent IP-rerouting capabilities, but they are unacceptably slow for some applications. Therefore, some fast protection techniques have been proposed in the IETF, profiting from the path-oriented nature of the MPLS technology.

MPLS protection

MPLS protection is based on pre-established disjoint backup Label Switched Paths (LSPs). Several recovery mechanisms are possible: Local Link Protection, Local Loop-Back and Path Protection ([huang], [makam]). And for all of them there are variants whether the backup paths are link or node disjoint from the working paths. Furthermore, restoration techniques are also employable in MPLS. See [colle1] and [colle2] for an elaboration on these topics. For the sake of brevity, we only consider *node disjoint* Path Protection: see Figure 3.

In **Path Protection (PP)**, for each working LSP, a pre-established disjoint backup LSP is set up, spanning the working LSP from ingress LSR to egress LSR. The ingress node is called the Protection Switch LSR (PSL), because this Label Switched Router (LSR) has to choose on which (working or backup) LSP to forward the packets. The egress node is called the Protection Merge LSR (PML), since it simply merges (*thus no "protection switching"*) both working and backup LSPs into the downstream part of the LSP. As long as the primary LSP is not failing, the ingress LSR forwards packets along this LSP. When the primary LSP is failing, the ingress LSR switches over to the backup LSP. The advantage of the global nature of PP is that only a single backup LSP is required per working LSP.

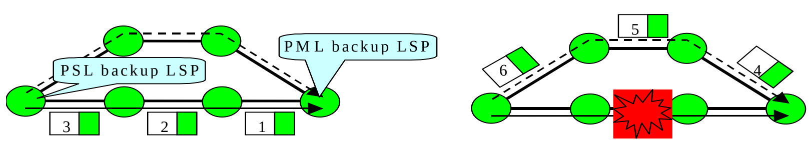


Figure 3: path protection, before and after protection switching

Some interesting features for providing PP in MPLS are :

- No resources have to be immediately allocated while setting up LSPs. An LSP only consumes resources when packets are forwarded through this LSP. Consequently, spare capacity can be shared and this results in a lower cost.
- An LSR can merge multiple incoming LSPs into a single outgoing LSP, if these LSPs have the same destination and their routes overlap downstream. So, the requirement for PP that the primary and the backup LSP can be merged in the downstream node, is no problem in MPLS.

MPλS protection

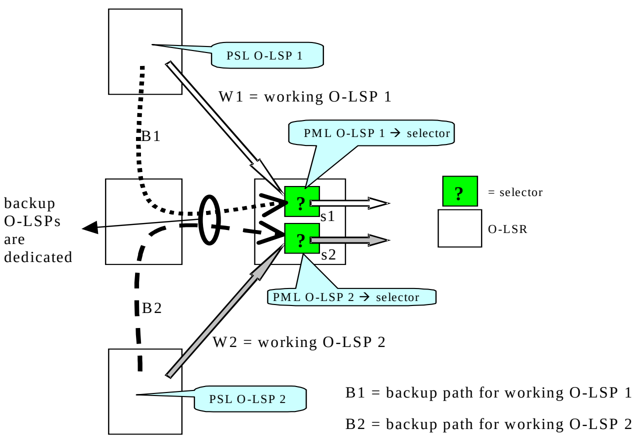
MPLS recovery techniques can also be adopted in an MPλS network.

Due to the physical meaning of a label in MPλS (being a wavelength) and the typical characteristics of optical networks, some powerful characteristics of MPLS recovery techniques can not be extended towards the optical layer.

In this context, the following issues concerning pre-established backup paths (backup O-LSPs) in MPλS networks are important (see Figure 4) :

Issue 1 – The number of wavelengths (optical labels) in a WDM point-to-point system is limited. This is the "capacity bottleneck" in optical networks. From the moment that a backup LSP is setup, it consumes a full wavelength (optical label), which is the capacity unit for optical networks and is a scarce resource. This leads to *dedicated* instead of shared (as in regular packet-based MPLS) protection. Without any improvements, two backup O-LSPs, protecting different failures, cannot share a wavelength.

Issue 2 – In current Optical Network Elements (ONEs), it is impossible to merge multiple O-LSPs into a single outgoing O-LSP. It is not possible to merge for example two 10 Gbps signals into a single 10 Gbps signal. A solution is to simulate the merging of the working and backup O-LSPs by using a *selector* *component*. The O-PML senses the incoming O-LSPs in order to decide by its own, which of both signals coming into the selector, has to be forwarded (the upstream O-PSL will always send the signal along one of both paths). The selector decides this according to which O-LSP is “inactive”. To indicate that an O-LSP is “inactive”, one could send no signal at all along that inactive O-LSP, allowing the selector to be implemented as a passive combiner, or one could use the framing overhead (e.g., SDH or Digital Wrapper (DW) framing) of the channel.



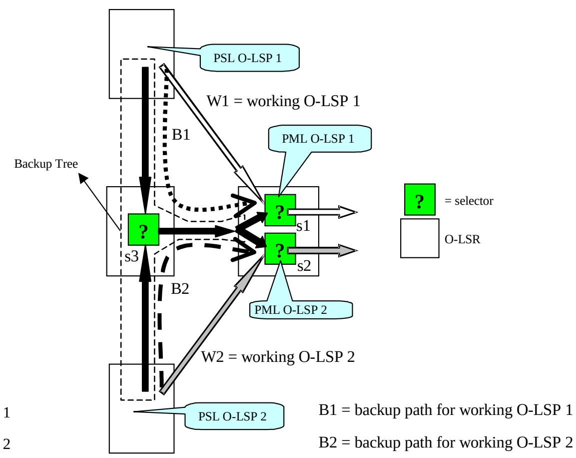


Figure 4 : MPλS protection based on pre-established backup paths, selectors are needed for merging and the protection is dedicated

Figure 5: a single-ended backup tree

We carried out some simulations to quantitatively analyze the dramatically increase in spare capacity cost as a result from Issue1 (when spare resources can't be shared). In Figure 6, the required spare capacity relative to the working capacity is depicted. This is done for the MPλS case as well as for the MPLS case. The working capacity cost is the same for both cases. Path protection was used as the recovery mechanism and single node failures where simulated. The values represent average simulation results over random demand matrices. The cost is the sum of the costs of the links. The latter are proportional with the link weight and the capacity that has to be carried by the link. This link capacity is (of course) the result of our dimensioning : we assumed shortest path routing for the working and backup LSPs. The severe impact of dedication in protection (as opposed to shared protection) is clear from Figure 6. Other simulation results can be found in [colle2].

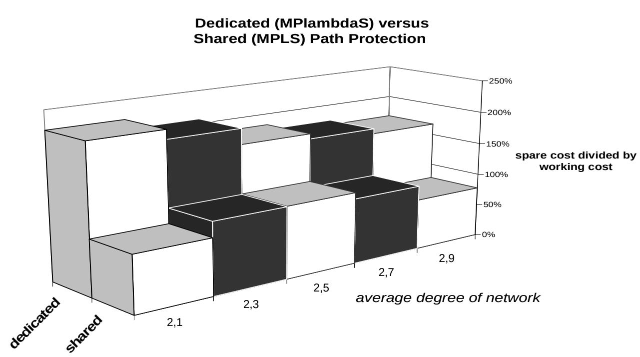


Figure 6: spare capacity resource comparison between MPλS (optical) and MPLS (electrical) Path Protection for several network variants

Additionally, Figure 6 shows that for shared PP the values bear no clear relationship with a lower nodal degree. While for dedicated PP a lower degree seems to a increase the cost for the protection paths, more than the cost for the working paths. The reason for this is that for a lower degree backup O-LSPs are long (relatively long to the working O-LSPs) and are overlapping a lot. Dedicated protection when using MPλS cannot profit from this overlap, although those backup O-LSPs protect against different failures. No capacity reuse (sharing) is possible.

Consequently, it is worthwhile to investigate potential improvements to cut down capacity costs. In the next section, we present such an improvement. It is obvious that a more advanced dedicated protection scheme will perform better than the original dedicated protection (=MPλS protection without improvements), but worse than shared protection (=MPLS protection).

Backup Tree concept

The backup tree concept presents some potential improvements for the protection schemes, in order to avoid the worst case dedicated protection (see section before).

Recall Figure 4: although the 2 backup O-LSPs partially overlap and protect against different single failures, the backup O-LSPs are dedicated and don't share capacity on their second link. But, if it is possible to simulate the merging of a working and a backup O-LSP in the PML with a selector (see Figure 4), then such a selector could also be put in use elsewhere and we could merge the overlapping parts of the 2 backup wavelengths (O-LSPs) into 1 wavelength.

This approach results in a structure that is called a (**Single-Ended) Backup Tree (SEBT)**. A Backup Tree is a set of wavelengths chosen on certain links (at most one per link) forming a directed tree towards a root node. This tree serves as a backup for multiple working lightpaths instead of providing a backup lightpath per individual working lightpath. The multiple working lightpaths that are to be protected by the backup tree, must be disjoint, so that single failure are correctly recovered from. A Backup Tree takes advantage of the fact that backup O-LSPs with the same destination (PML) and which overlap completely from an intermediate O-LSR x to the PML and which protect different lines or equipment (links or nodes), can be merged into a single outgoing path from O-LSR x until the PML. A SEBT only protects working lightpaths to the same PML node (the destination node). So all working lightpaths are originating from a leaf node in the tree and are terminating in the root of the tree. This restriction can be removed, resulting in an extension that is called a Multi-Ended Backup Tree (MEBT). See [colle2] for details on this.

Now, look at Figure 5 and compare with Figure 4 to notice the backup tree that will replace the original backup paths B1 and B2. The SEBT protects two working paths W1 and W2. Without the SEBT improvement, the backup path B1 is protecting W1: B1 is a disjoint path with W1 originating from the same source and going to the same destination. Selector s1 normally selects the (optical) signal coming from W1, but chooses B1 when there is a failure along W1. The same applies to W2 and B2, mutatis mutandis. Using the SEBT improvement, B1 and B2 are *replaced* by the backup tree structure that includes the selector s3 and which splits out in the destination node to both selectors s1 and s2. Normally, s1 and s2 are selecting the working signals (W1 and W2) and the selection of s3 doesn't matter. If W2 fails, PSL for O-LSP 2 will switch over to the backup tree. Selector s3 will select the signal coming from O-LSP 2, because this is the only (valid) signal. Selector s2 will select the signal coming from the backup tree, because there is no signal from W2. From this we can conclude that both W2 and W1 have a protection path. But now, in comparison with Figure 4, only three backup wavelengths are needed as a result of the reduction in spare capacity by share the wavelength on the second link of the original backup paths (B1 and B2).

Remark that it is clarifying to see a backup tree as a "shared representation" of a collection of overlapping backup paths (in Figure 5 for example : B1 and B2). Also, note that the tree-structure of the SEBT assures that backup O-LSPs do not divert. If this would be the case, then at the O-LSR where the backup O-LSPs divert, one would not be able to decide along which path to forward the backup signal. The (selector in the) O-LSR can't take the appropriate decision, since it only knows that it has to forward the backup signal, but not where the backup signal will be sent to, as he doesn't know to what primary O-LSP the backup signal belongs.

Finally, with every (destination) node in an OTN, there is a *set* of backup trees associated. Each backup tree will protect a *set* of disjoint working paths.

Finding optimal backup trees: problem analysis

The next question is how to set-up the working paths and the backup trees in a way that consumes the minimal amount of capacity resources (wavelengths). Let's look how to optimally provide *path protection using a SEBT (SEBT for PP problem)*.

The SEBT for PP problem is in essence : "Working paths should obviously fulfil the working demand and should be placed so that as much of them are mutually disjoint. A backup path should be disjoint with the working path it is protecting. Backup paths of disjoint working paths should be placed so that as much of them are overlapping as much a possible. Backup paths of overlapping working paths bear no relationship. "

The problem is mathematically formulated in Figure 7 and Figure 8. Not all equations are in an ILP form, but it is possible to go to an ILP formulation by introducing new binary variables and/or splitting up into more equations.

- Given :
  - a directed graph $G = (V, A)$, *with node-connectivity* ≥ 2
  - a cost function on the arcs : $c(a) \geq 0, \forall a \in A$ (the cost of using a wavelength on that arc)
  - the demand $d(i,j,k), \forall i \in V, \forall j \in V, \forall k \in \{1, 2, \ldots, M\}$ (from node i to node j), with

$$d(i,j,k) = \begin{cases} 1 & \text{if at least } k \text{ wavelengths paths are demanded,} \\ 0 & \text{if less than } k \text{ wavelength paths are demanded.} \end{cases}$$

- Definition : $\forall v \in V : \begin{cases} \mathcal{I}(v) & = \text{all arcs leaving node } v \\ \mathcal{I}'(v) & = \text{all arcs arriving at node } v \end{cases}$

- Definitions :

$$C = \{(i,j,k) \mid i \in V, j \in V, k \in \{1, 2, \ldots, M\}, d(i,j,k) = 1\}$$
$$= \text{set of the existing wavelength connections}$$

$$\forall j \in V : C(j) = \{(i,k) \mid i \in V, k \in \{1, 2, \ldots, M\}, (i,j,k) \in C\}$$
$$= \text{set of the wavelength connections towards the node } j$$

$$\forall j \in V : \mathrm{P}(j) = \left\{ s_1^j, s_2^j, \ldots, s_{|C(j)|}^j \right\}$$

= the collection of sets representing a partition of the connections towards the node j (each such set s will contain a series of connections that will be protected by the same backup tree)

- Variables : $\forall j \in V$:
  - $\forall (i,k) \in C(j), \forall a \in A : f(i,j,k,a)$: representing the flow over arc a that is part of the working path for the wavelength connection (i,j,k) (binary variable)
  - $\forall (i,k) \in C(j), \forall s \in \mathrm{P}(j) : \delta(i,j,k,s)$: binary variable, describing the partitioning of the connections towards the node j ($\delta(i,j,k,s) = 1$ means that connection (i,j,k) belongs to set $s \in \mathrm{P}(j)$)
  - $\forall (i,k) \in C(j), \forall s \in \mathrm{P}(j), \forall a' \in A : g(i,j,k,a')$: representing the flow over arc a' that is part of the backup path that forms the protection path for the working connection (i,j,k) (binary variable)
  - $\forall s \in \mathrm{P}(j), \forall a' \in A : v(j,s,a')$: binary variable, choosing the arcs a' that belong to the backup tree associated with set s ($v(j,s,a') = 1$ means that arc a' is selected to belong to backup tree s)

- Objective : minimize the cost of the working paths and the backup trees

$$\min \sum_{j \in V} \sum_{a' \in A} \left\langle \sum_{s \in \mathrm{P}(j)} v(j,s,a') + \sum_{(i,k) \in C(j)} f(i,j,k,a') \right\rangle \cdot c(a') \qquad (1)$$

Figure 7: first part of mathematical formulation

This formulation shows that the problem can be split up into apart subproblems : one subproblem for each destination node j. Furthermore, the idea is to partition the set of connections towards a destination node j, into a series of sets. For each set, it must be possible to set-up the working paths in such a way that a backup tree, incorporating the backups for these working paths, can be constructed. The general objective thereby is to minimise the total cost of using wavelengths in the OTN.

From the formulation, it is obvious that a brute force ILP method won't be feasible : the number of binary variables is far too large for realistic network sizes. In order to tackle the problem, we will start to investigate subproblems like "given a fixed placement of working (and even backup paths) find the optimal partitioning of the working paths". From the mathematical formulation, it is already clear that sets and partitions will be encountered. It will be useful to look for efficient problem-specific ways to reduce exhaustive enumeration of these combinatorial structures. We will investigate whether an optimal algorithm based on Integer Linear Programming seems feasible, or whether a heuristic construction or search method will be more suitable.

Of course, the problem has many extensions : extending to the less restrictive MEBT structure, evaluating variants concerning other protection types,… Research results covering these topics will be shown at the conference.

Constraints for the flows that make up the working connections :

$$\forall (i,k) \in C(j), \forall v \in V : \sum_{a \in \mathcal{I}(v)} f(i,j,k,a) - \sum_{a \in \mathcal{I}'(v)} f(i,j,k,a) = \begin{cases} 1 & \text{if } v = i, \\ -1 & \text{if } v = j, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Partitioning the different connections towards a node j into the sets $s \in \mathrm{P}(j)$:

$$\forall (i,k) \in C(j) : \sum_{s \in \mathrm{P}(j)} \delta(i,j,k,s) = 1 \quad (3)$$

Only connections that are disjoint (in their intermediate nodes) are allowed to settle in the same set $s \in \mathrm{P}(j)$:

$$\forall s \in \mathrm{P}(j), \forall v' \in V \setminus \{j\} : \sum_{(i,k) \in C(j)} \delta(i,j,k,s) \left(\sum_{a \in \mathcal{I}'(v')} f(i,j,k,a) \right) \leq 1 \quad (4)$$

Constraints for the flows that make up the backup paths:

$$\forall (i,k) \in C(j), \forall v' \in V : \sum_{a' \in \mathcal{I}(v')} g(i,j,k,a') - \sum_{a' \in \mathcal{I}'(v')} g(i,j,k,a') = \begin{cases} 1 & \text{if } v' = i, \\ -1 & \text{if } v' = j, \\ 0 & \text{otherwise .} \end{cases} \quad (5)$$

A "backup path" must be node-disjoint with the "working path" it protects :

$$\forall (i,k) \in C(j), \forall v' \in V \setminus \{j\} : \sum_{a \in \mathcal{I}'(v')} (f(i,j,k,a) + g(i,j,k,a)) \leq 1 \quad (6)$$

Derive the arcs of the backup tree s : while a connection (i,j,k) is selected to be protected by the backup tree s, the backup tree will contain arc a' when the backup path for this connection indeed flows through this arc a':

$$\forall s \in \mathrm{P}(j), \forall (i,k) \in C(j), \forall a' \in A : \delta(i,j,k,s) = 1 \Rightarrow g(i,j,k,a') \leq v(j,s,a') \quad (7)$$

The destination node j in a backup tree $s \in \mathrm{P}(j)$ can only have one incoming backup tree arc (=choice in definition) :

$$\forall s \in \mathrm{P}(j) : \sum_{a' \in \mathcal{I}'(j)} v(j,s,a') \leq 1 \quad (8)$$

Restrict the backup tree arcs to a tree structure :

$$\forall s \in \mathrm{P}(j), \forall v' \in V : \sum_{a' \in \mathcal{I}(v')} v(j,s,a') \leq 1 \quad (9)$$

Figure 8: 2nd part of mathematical formulation, constraints *for all nodes $j \in V$*

Conclusions

In this paper, a novel concept to provide protection capacity for MPλS recovery mechanisms is introduced : a Backup Tree. This technique should allow for a more efficient resource usage than the fully dedicated MPλS protection solution. The problem of determining the placement of these tree structures that is optimal with respect to the needed capacity, is presented in an ILP-like formulation (focussing on the case of Path Protection). At the conference, optimal or heuristic algorithms will be presented to estimate the capacity savings that can be expected by the use of Backup Trees. Quantitative results for various network situations will be reported.

Acknowledgements

Part of this work has been supported by the European Commission through the IST-project LION and by the Flemish Government through the IWT-project ITA/980272/INTEC and by the Fund for Scientific Research - Flanders (Belgium) (the first and fourth author are research assistants of the F.W.O.-Vl.).

References

[colle1] Didier Colle, et. al., "MPLS Recovery Mechanisms for IP-over-WDM networks", special issue on 'IP over WDM and Optical Packet Switching' of 'Photonic Network Communications Magazine', Vol.3, No.1, January 2001, to be published.

[colle2] Didier Colle, et. al., "Porting MPLS-recovery techniques to the MPλS paradigm", accepted for publication in special issue on 'Protection and Survivability in Optical Networks' of 'Optical Networks Magazine' (2001)

[ghani] Nasir Ghani, "Lambda-Labeling: A Framework for IP-over-WDM using MPLS", Optical Networks Magazine, Vol. 1, No. 2, April 2000, pp. 45-58.

[manzalini] Antonio Manzalini, "Milestones for the Evolution toward an Integrated Optical Transport Network", Optical Networks Magazine, Vol. 1, No. 1, January 2000, pp. 29-34.

[makam] Makam, et. al., "Protection/Restoration of MPLS networks", work in progress, Internet-Draft, October 1999, http://search.ietf.org/internet-drafts.

[huang] Changcheng Huang, et. al., "A Path Protection/Restoration Mechanism for MPLS Networks", work in progress, Internet-Draft, March 2000, http://search.ietf.org/internet-drafts.

[awduche] Daniel Awduche, et al., "Multi-Protocol Lambda Switching : Combining MPLS Traffic Engineering Control With Optical Crossconnects", Internet Draft, work in progress, 1999, http://search.ietf.org/internet-drafts.