

Informed Dynamic Shared Path Protection

David Elie-Dit-Cosaque, Maher Ali and Lubo Tancevski

Alcatel Research and Innovation, M/S CTO2, Richardson, TX 75081

Tel: (972) 996-3936, Fax: (972) 996-5174, Email: David.Elie_Dit_Cosaque@alcatel.com

Abstract: We propose a new concept for distributed mesh protection where the information about protection resources at every link is shared throughout the network, resulting in substantial savings in terms of allocated protection capacity.

©2000 Optical Society of America

OCIS codes: (1 060.4250) Networks.

Introduction

A generalized optical network consists of multiple optical cross-connects OXC that are interconnected via bi-directional WDM links in a general topology, often referred to as intelligent optical mesh network. A major aspect of this intelligence is the ability to provide fast provisioning and restoration. To achieve this, there is a need for a unified control plane responsible for running routing and traffic engineering protocols, achieved, for example, through the introduction of GMPLS. The control plane queries and modifies the wavelength-switching table of the cross-connect; floods the reservation tables of the cross-connect (through OSPF or IS-IS) in the network; is responsible for connection set-up/tear-down through CR-LDP (or RSVP-TE). A mesh-restored lightpath in an optical network is allocated a pair of link-disjoint paths, where one of the path is the primary or working path and the other is backup or protection path that is activated only in case of failure. Each link in the primary path has dedicated capacity allocated to a connection. The protection path can also have dedicated capacity (1+1 restored lightpath), however that results in inefficient use of network capacity. In contrast, in shared or 1:N restored lightpath, the protection capacity is shared with the backup path for other restored connections, hence resulting in improved utilization of the network resources [2]. Shared route computation can proceed through centralized network management system that results in good network utilization but does not scale well with network size, can be a single point of failure and is rather slow for dynamic applications. An alternative is the distributed computation with a use of a database at every OXC, where a decision on sharing is being done locally on a link-by-link basis. The distributed approach provides better scalability, resilience and speed, but results in sub-optimal sharing and hence utilization of network capacity. In this paper we propose a novel distributed approach, called Informed Dynamic Shared Path Protection (IDSPP), where information about the shared resources on every link is distributed throughout the network using extensions to OSPF (or IS-IS), resulting in much better sharing. Simulation results indicate substantial improvement over the uninformed (1:N) approach.

Protection algorithm description

The basic idea behind the concept is to maintain a protection-sharing table at every OXC that keeps track of the resources allocated for protection. The information contained in the table can range from partial (where only the number of paths protected on a particular link is maintained) to complete (where information on every path is stored). A general example of such table is depicted in Fig. 1. For each wavelength set aside for protection there is an entry listing the protected working paths. For instance, Wavelength 2 is set aside for protecting working paths {Node1, Node8, Node2} and {Node1, Node4}. Next, the content of the protection-sharing table is flooded. The OSPF (or IS-IS) protocol can be used for the distributed network database synchronization. It needs to be emphasized that in order not to escalate the bandwidth in the OSPF, only recent modifications in the table are distributed, and not the entire table. This results in a global knowledge of both network and protection topology (in terms of sharing capability as well as wavelength allocation) allowing both working and protection route computation at the source.

Elements of the computation are as follows: The path to the destination is the shortest path in terms of cost;

Wavelength	Paths sharing this link
1	{1, 2, 3, 7}
2	{1, 8, 2}, {1, 4}
...	

The length of the link and its capability to share wavelength for protection defines its cost. The calculation of the cost for each link is done as following: if the link can be shared for protection, the resulting cost will only be the length of the link, and if not, the resulting cost will be the length of the link multiplied by a *penalty* coefficient. Therefore the algorithm uses first the segments where

Figure 1: Protection sharing table.

sharing is possible, optimizing the reuse protection resources. Next, the source node sends a connection request message to the destination in order to reserve resources along the computed path. However, the network may reject this request. We do not introduce any retry strategies. In this case, both the working and the protection lighthpath fail.

An improved version of Label distribution protocol (LDP) that handles optical connections is proposed. The protocol provides the capability to setup and tear down a lightpath through exchanging messages to reserve or release the optical resources. A message can either be SETUP, ACCEPT, REALEASE, TERMINATION, FAILURE NOTIFICATION or ACTIVATION. The SETUP procedure consists of the following steps: the shortest path to the destination is calculated by the source. Next, the wavelength allocation process is carried out. The first fit algorithm selects a wavelength on the first segment of the working path. The algorithm tries to reserve the same wavelength in the following segments. If it is not successful, a wavelength conversion is requested. The Connection SETUP message is sent by the source along the calculated path. If calculated resources are available in every segments of the path, the destination receives the SETUP message and acknowledges it positively by sending an ACCEPT message back to the source. Finally, the data is sent by the source. If contention occurs, a RELEASE message is immediately sent back to the source. The destination node does not receive notification of failure. In order to terminate a connection, a TERMINATION message is sent to the destination along the reserved path. Each node receiving the message releases the resources associated with the terminated connection. The protection path is established just after the calculation of the working path. The algorithm calculates the path in order to maximize the sharing of costly resources like converters and ports as described previously. The connection setup is almost the same for protection. The only difference is that the sharing table is updated when the SETUP message is received: the working path is added to the list of paths sharing this wavelength for protection. If the reservation succeeds, the protection path is added to the *protection-sharing* table in the source node.

When a link failure occurs an alarm is generated and a NOTIFICATION message is sent to the source. Once the message is received, the source node makes a lookup on its *protection-sharing* table. Given a triplet (source, destination, working connection ID), the lookup operation returns the protection path and the resources (wavelength, regeneration) to use on each portion of the backup path. Because the path is reserved, only the switching time should be considered in the restoration process. In other words, we do not wait for the path to be completely established to forward the data. Indeed, in case of a single failure, contention on the shared resources does not happen.

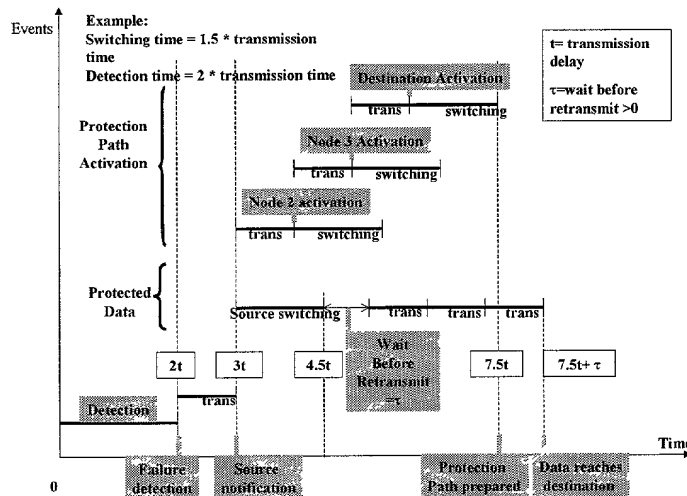


Figure 2: Steps for protection switching.

Fig. 2 shows the sequence of events occurring during the restoration process. The protection switching time is the following for the example in Fig 2:

Protection Switching Time = 1 * Switching time (10ms) + detection time (12ms) + 1 transmission delay (6ms) + wait-before-retransmit (5ms) ≈ 33 ms

Results

We consider the Italian network in our simulation. The traffic consists of 500 connection requests. One connection is defined by 2 parameters: a) start time, b) duration. The start time and the duration are uniformly distributed on the interval [0..200 seconds]. In order to quantify the gain due to dynamic shared protection, we compare two distributed approaches: 1:N shared path protection without protection topology

knowledge (uninformed) and 1:N shared path protection algorithm that is topology aware (informed). Several simulations were run with different traffic loads. We progressively increase the number of connection from 0 to 500. Fig 3 shows the efficiency in terms of wavelength mileage of the two protection schemes (uninformed 1:N and informed IDSP with penalty 5) and the working mileage as a function of the established connections. As evident the informed protocol always requires less wavelengths since the sharing of the resources is much better. For example, for 400 established connections, the informed protocol is 35% more efficient in terms of wavelength mileage compared to uninformed (18,000 miles savings). When the number of connections increases, the improvement also increases due to more resources being available for sharing - when a protection path is

established, the probability of link sharing is increased. In Fig 4, we observe that the informed algorithm generates longer paths than the uninformed. The explanation is that the uninformed algorithm runs the shortest path algorithm to compute the backup route. When possible, it shares the link by doing a decision on a hop-by-hop basis during the

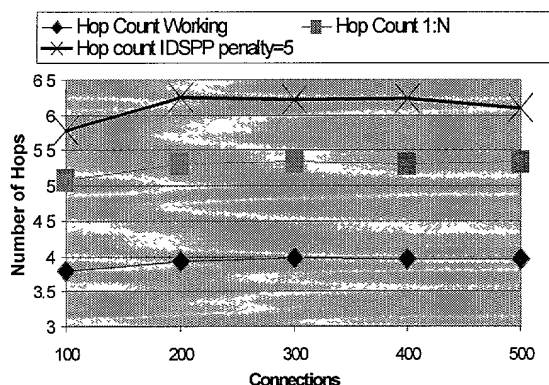


Fig. 4: Path Hop count for informed and uninformed algorithms (penalty coefficient=5).

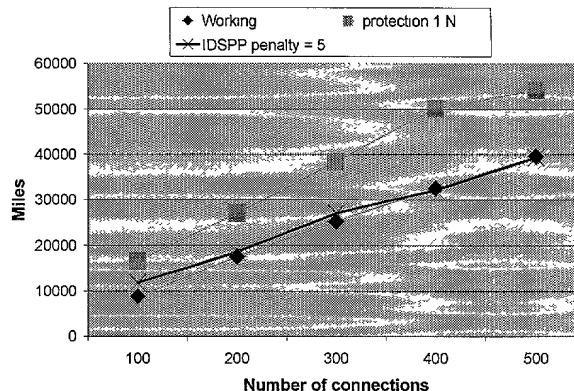


Fig. 3: Wavelength Mileage for informed and uninformed algorithms (penalty coefficient=5).

wavelength allocation phase. In the informed algorithm, on the other hand, a priority is given to increase the reuse of shared resources when computing the path in the source node through the introduction of a penalty function. This effect increases the protection mileage of the informed scheme, but that is more than offset by the better sharing of resources, as evident from Fig. 3. Fig 5. shows the difference (calculated as number of shared paths per link for the informed minus the number of shared paths for the uninformed) in terms of sharing between the informed and the uninformed protection schemes. The difference is calculated from the average values of the number of paths shared per link over time for each particular link. The informed protocol shares less links (69 versus 72 for the uninformed) as is expected due to better sharing. However, the number of protected paths over the links that are shared is increased, indeed indicating better sharing. This is evident from Fig. 5 where the positive difference outweighs the negative side of the graph. The maximum average number of shared backup paths (obtained over all links) for the informed approach was found to be 22 compared to 15 for the uninformed approach.

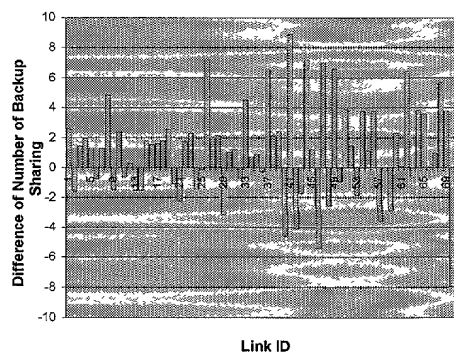


Fig. 5: Comparison between the sharing of informed and uninformed protection.

Conclusion

The efficiency of Distributed 1:N mesh restoration can be enhanced substantially by providing information about the protection topology. This information can be disseminated using extensions to OSPF. In order to keep the bandwidth of the distributed information low, only recent modifications in the protection tables need be distributed. Calculations indicate substantial savings (up to 35%) in terms of protection wavelength mileage due to much better sharing of backup resources. In addition, the scheme allows for different algorithms when choosing paths that share a link, thus potentially enabling protection in case of multiple failures.

References

- [1] R. Doverspike et al., "Challenges for MPLS in Optical Network Restoration," *IEEE Communications Magazine*, Feb. 2001.
- [2] S. Ramamurthy et al., "Survivable WDM Mesh Networks, Part II – Restoration", IEEE, 1999.

QoS –based Routing in Optical Networks

Maher Ali, Girish Chiruvolu, David Elie-Dit-Cosaque and Lubo Tancevski

Alcatel Research and Innovation, M/S CTO2, Richardson, TX 75081
Tel (972) 996-3571, Fax (972) 996-5174, Email Maher.Ali@alcatel.com

Abstract: We address issues related to quality-of-service (QoS) differentiation in the optical layer. Numerical examples are presented showing the effect of providing differentiated connections on both the cost and signal quality.

©2000 Optical Society of America

OCIS codes: (1 060.4250) Networks.

Introduction: The intelligent optical networks paradigm, where optical cross-connects are interconnected with WDM links, requires tight integration between the IP and the optical layer such that: the non-homogeneity of the optical layer is taken into account during the routing process; there is mapping of the QoS requirements down to the optical layer. This effectively enables the introduction of optical QoS differentiation. Optical QoS can be classified into two categories: 1) *quantitative measures* of the optical signal itself such as bit-error-rate (BER), power, etc. [1]; and 2) *functional requirements* such as lightpath routing stability and survivability [2]. Both categories allow assigning differentiated guarantees for different classes of service. In this paper, we address both aspects of optical QoS and through numerical examples we show that providing differentiation in terms of signal quality, routing stability and survivability leads to better utilization of available network resources.

Quantitative Measures: Recently, there has been some activities concentrating on addressing the impact of transmission impairments on the routing process and the ensuing network optimization [1]. Transmission impairments can be classified into two categories: *linear* and *nonlinear* impairments. Example of linear impairments is the PMD effect, more apparent in higher data rates (10Gb/s and above) and believed to be the dominant impairment effect for 40Gb/s networks. The interesting problem that arises due to transmission impairments is the heterogeneity of the fiber links. Old fiber links can inflict more impairments. These impairments have the side effects of increasing the regeneration budget as well as the overall BER. Such old fibers can have average PMD coefficient ranging from 2 to 0.5 ps/√km. New and improved fibers lessen, to some extent, the PMD effect and have PMD coefficient around 0.1 ps/√km. The non-homogeneous optical resources mandates that the control plane routing process takes the transmission quality of the optical fibers into considerations during the *routing* phase. Information about the type of fiber and the transmission impairments can be distributed using extensions to OSPF or IS-IS, and that information can be taken into account when performing routing such that the shortest path metrics is changed from least number of hops to minimum induced noise instead of least number of hops [1].

One can envision different classes of traffic offered in a wavelength-routed network. These classes can vary in their maximum BER requirement, survivability guarantees, etc. The optical network must be able to differentiate between these classes of traffic allowing for proper connection admission as well as better utilization of the heterogeneous optical resources. From the quantitative measures perspective, different classes will have different demands in terms of signal quality and hence will be routed differently – higher class traffic will be routed through links with better transmission quality. A typical example will assume separating the traffic in terms of OC-192 and OC-768 and routing them separately. However, whereas no blocking is allowed in OC-192-based applications [1], substantial blocking can occur in case of traffic routed as OC-768, as indicated in Fig. 1. In this scenario, links that have 0.5 ps/√km PMD coefficient are eliminated from the route-selection scheme; thus dramatically increasing the blocking probability. To overcome this increase in 40Gb/s rejection, routing algorithms and signaling protocols must be extended so that a 40Gb/s signal can be *split* into four 10Gb/s signals when passing through an old fiber (intermediate splitting) or alternatively, the entire traffic must be split from source to destination into 10Gb/s and 40Gb/s connections which are routed separately (source splitting). Note that source splitting effectively accomplishes QoS differentiation between 10 and 40Gb/s traffic. To investigate the source splitting scheme, simulations were performed for a typical network topology; fibers are selected randomly to be new or old fibers up to a percentage; the information about the PMD coefficient for every link is distributed using OSPF; routing metrics is based on least accumulated noise. Fig. 2 shows the advantage of source splitting on the network utilization. In the figure, the ratio of the bandwidth successfully carried, using the choice of route splitting to that of no route splitting, is plotted on the y-axis as a function of the percentage of “good” fiber in the network. As expected, when the old fiber dominates the good fiber, the improvement is drastic due to the fact that most of the available links are being eliminated for establishing OC-786 connections due to unacceptably high PMD.

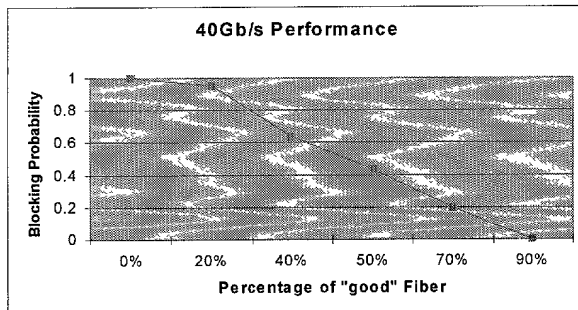


Fig. 1. 40Gb/s blocking performance.

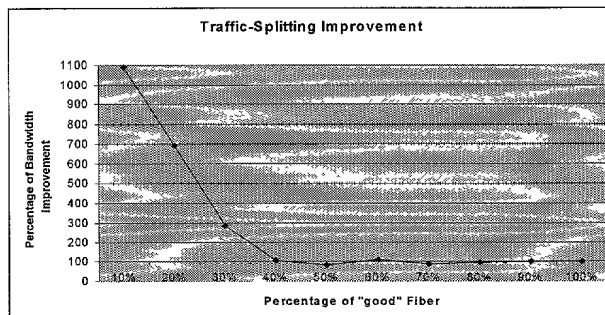


Fig. 2. Improvement under source splitting.

Functional requirements - routing stability: Integrated traffic engineering (TE) for IP/WDM is an active area of research and efforts are currently ongoing in the standards bodies such as IETF to devise appropriate tools/mechanisms over which a TE can be realized. The integrated TE should address the issues of dynamic establishment of the lightpaths in an optimal manner so as to accommodate maximum amount of the user's traffic on demand. Routing stability is a key issue in such an integrated TE and in this subsection we briefly discuss an approach towards the integrated TE, referred to as RITE.

Routing stability for individual traffic trunks is critical, since many relevant QoS measures such as delay, jitter, and throughput can be degraded. For example, TCP applications that do not have large buffers for re-ordering of the out-of-sequence packets may trigger retransmission of many packets. Delay and jitter are the other measures that are critical to real-time applications that also get affected as a result of rerouting of traffic trunks. In the case of non-packet-based TDM signals, rerouting may translate into disruption of the ongoing signals. TE algorithms should address this issue in the context of dynamic wavelength assignment and traffic mapping in order to make optimal use of the optical resources.

The RITE scheme is geared towards an MPLS/Optical network, where in the border nodes (ingress/egress) are the MPLS routers that are connected by an optical network that imposes wavelength continuity constraint across the end points (MPLS routers). Light channels (LCs) are established on demand depending on the traffic loads at the MPLS routers. The goal is to increase the utilization of the existing/established light channels (LCs) carrying maximum traffic, while conservatively establish new LCs and minimize the rerouting frequency of traffic trunks during their lifetime (period between the time of establishment to the tear down of the traffic trunk at the ingress routers). The LCs that are underutilized are torn down, thereby freeing up the relevant optical resources (wavelengths) and LCs are established based on the incoming traffic demands. The traffic is classified as high priority (HP) traffic that requires higher routing stability and the low priority (LP) traffic that can tolerate QoS degradation due to limited rerouting but subject to a limit on the frequency of rerouting. The HP traffic trunks are mapped on to the direct LCs (/LSPs) that connect the given pair of ingress and egress MPLS nodes and are rerouted only in the event of LC tear down due to poor traffic utilization. The LP traffic trunks are mapped on to the direct LCs if available. Otherwise, they are mapped on to LSPs that span multiple LCs with appropriate optical-to-electronic (O-E, vice-versa) conversions at the MPLS routers serving as the intermediate hops (multi-hop LSP). The

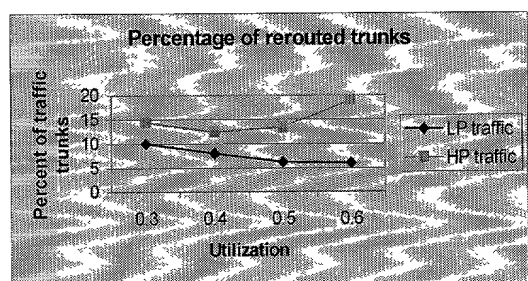


Fig. 3: Performance of RITE

tear down notification. Thus, each LC is associated with a time-to-live (TTL) timer and is updated accordingly to reflect the utilization states. These LSAs are noted and the appropriate tables are updated at the MPLS/Optical

O-E-O conversions at the intermediate routers may introduce packet delays for the traffic mapped in to the multi-hop LSPs. Each such LP traffic trunk is associated with a rerouting timer that is set at the time of rerouting, so as to prevent another rerouting until the timer expires. The RITE's policy is to accommodate HP traffic on the existing direct LCs and any existing LP traffic trunks are rerouted freeing up the bandwidth for HP traffic. Each of the LCs is associated with four types of link state advertisements (LSAs), viz., a) LC establishment, b) LC on the verge of over utilization wherein the traffic reaching an upper watermark, c) LC utilization falling below a lower watermark triggering a tear down (release) of the LC, d) LC

nodes. The First Fit (FF) scheme is employed to establish the LCs across the optical network to choose the wavelengths. The HP/LP traffic bandwidth requests are of fixed units and their active time intervals are drawn from an exponential distribution. Fig. 3 demonstrates the effective routing stability provided for the HP traffic, while the rerouting frequency for LP traffic is bounded within 20% during their lifetimes which are exponentially distributed. When the utilization of the LCs is low thereby triggering the tear down of the respective LCs some of the HP traffic is also rerouted and mapped on to multi-hop LSPs. For the LP traffic, as the utilization factor increases, the probability of rerouting of the LP trunks making way for HP trunks of direct LCs, increases. This demonstrates the feasibility of achieving routing stability under the above mentioned criteria and as a result, the QoS measures such as throughput, delay, jitter can be improved and the service signal disruption for the ongoing sessions can be minimized.

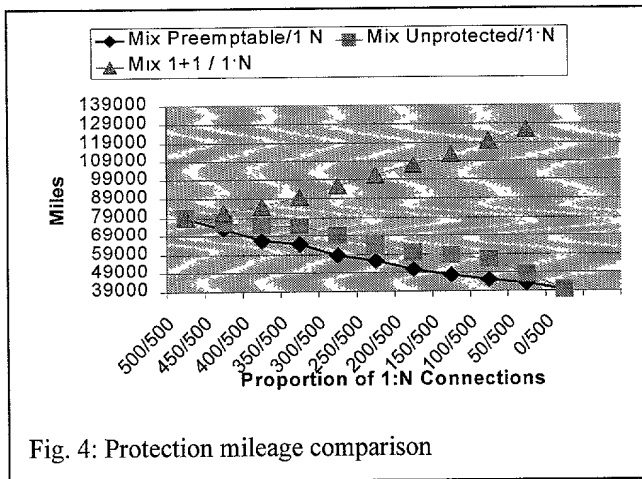


Fig. 4: Protection mileage comparison

mileage produced by a 500 connection requests. The connections start and duration are uniformly distributed over the interval [1, 200s]. We consider a typical network with sufficient link resources in order to avoid rejection. The input connections can implement gold class, preemptable class, or unprotected. The preemptable connections are assigned using the shortest path algorithm and may incorporate links shared for 1:N protection. The unprotected working class is introduced in order to make a comparison with preemptable class and, in difference to it, does not include links set aside for 1:N protection. From Fig. 4, we observe significant savings in terms of protection mileage, as expected, when we increase the proportion of preemptable class versus 1:N class. Replacing 150 1:N connections with preemptable ones (350/500 point) provides 12600 miles savings. On the other hand, replacing 150 1:N connections with unprotected ones saves only 4600 wavelength miles. This can be explained on the basis that preemptable connections can allocate either new resources or reserved inactive resources. The inactive resources are reserved by the gold class connection in order to guarantee backup in case of failure of the working path. As a result, if a failure occurs, preemptable connections that use portion of the backup path are terminated. The fine-tuning of the preemptable traffic allocation algorithm can increase the use of inactive backup paths and, as a consequence, reduce the mileage. We observe substantial savings when comparing 1:N and 1+1 cases - for example, 1+1 class uses 39% more resources than the 1:N case for the 350/500 mixture.

Conclusions: In this paper two aspects of optical QoS, namely the quantitative measures of the optical signal and the functional requirements such as routing stability and protection were discussed. Transmission impairments-based routing of 40Gb/s connections was investigated and substantial improvement in bandwidth utilization was demonstrated when traffic was split at the source into 10 and 40Gb/s connections, that were routed separately based on the quality of the available transmission links. The feasibility of differentiation based on routing stability in integrated traffic engineering was demonstrated using the RITE scheme with the preliminary results indicating substantial potential for better utilization of existing lightpaths. Class-based protection and restoration has been investigated and numerical examples were provided that show how different classes of traffic could be routed in a cost-effective way requiring less protection bandwidth.

References

- [1] M. Ali, et al. "Network Optimization with Transmission Impairments-Based Routing," Proc. ECOC'01, Amsterdam, October 2001.
- [2] N. Golmie et al. "A differentiated optical services model for WDM networks," IEEE Communications Magazine, Volume: 38 Issue: 2, pp. 68-73, Feb. 2000.

Functional requirements - protection/restoration:

One of the most important functional measures of optical QoS is the survivability requirement of the lightpath. It can be envisioned that there are at least three classes of traffic demands: a) *Platinum*, b) *Gold*, and c) *Bronze*. For the platinum traffic, a 1+1 protection is provided such that, two live signals at the receiver are available to choose from. The gold service can provide a restoration guarantee by reserving backup resources and activating these resources (in a distributed protocol) for the lightpath when failure occurs on the primary path. The reserved resources can be shared among different backup paths depending on the fault model (1:N). The bronze service can be preemptable or unprotected traffic. Fig. 4 shows the protection λ