Enhancements to Multi-Protocol Lambda Switching (MPλS) to Accommodate Transmission Impairments

Maher Ali, David Elie-Dit-Cosaque, Lubo Tancevski
Alcatel Corporate Research Center
1201 E. Campbell Road
M/S CTO2
Richardson, TX 75081
U.S.A.

Abstract

We consider networks consisting of optical cross-connects interconnected via fiber optic links and controlled by IP routes. Optical connections are provisioned dynamically and on demand as the need for them arises due to changes in the traffic. We discuss the impact of linear effects such as *Polarization Mode Dispersion* (PMD) on the design of the control plane. Constrained-routing algorithms as well as signaling protocols are developed which allow for proper label (lambda) and regeneration requests distribution. Both the architecture and the electronic equipment cost are validated, through simulation, using a new WDM-aware version of NS-2 simulator. Simulation results reveal that: 1) the feasibility of the lightpath is greatly affected by the PMD factor, especially for bit-rates exceeding 5Gb/s, and 2) significant improvement (in terms of cost) can be achieved through intelligent route selection schemes which take into consideration physical characteristics of fiber links. These cost savings are more apparent in networks with non-homogeneous fiber quality.

I. Introduction

Optical networking is believed to be the answer to the rapid increase in bandwidth demands and the recent interests in providing Quality of Service (QoS) to the best-effort Internet Protocol (IP) [1]. The bandwidth problem can be solved in a cost-effective way by using wavelength-division multiplexing (WDM) [2] which allows for the transmission of many, relatively-independent, channels on the same fiber. Managing this huge bandwidth is turning out to be a challenging problem. Multi-protocol lambda switching (MPλS) promises intelligent usage of optical resources. MPλS builds on top of the existing traffic-engineering tools found in Multi-Protocol Label Switching (MPLS) [3] to provide a common platform for managing both the IP and the optical layers. In addition, it adds the capabilities for fast provisioning of wavelengths, λ-bands, and even fibers.

In this paper, we show how such a common management scheme can be deployed in optical networks where transmission impairments factors are non-negligible. This paper is organized as follows. This section continues by illustrating the environment we are considering. In Section II, a detailed treatment of the signaling protocol is presented. The traffic, tool, and the performance analysis of this enhanced protocol are discussed in Section III. Finally, we provide some conclusions in Section IV.

A. Network and Node Architectures

There are two basic network models for IP Over WDM: *overlay* and *peer* models. In the overly model, the IP edge router acts as a client to the optical network which provides a high-capacity on-demand pipes of bandwidth. The IP layer is unaware of the topology of the optical layer. Both layers are managed by two separate control planes. In the peer model, however, a single instance of the control plane is utilized to control both layers. An edge router has full knowledge of optical resources (e.g., OXCs, wavelengths, etc.) allowing more bandwidth-efficient usage of the scarce optical resources.

We consider networks composed of nodes (see Fig. 1) interconnected using optical fiber. Every two adjacent nodes have at least two fibers between them, one in each direction. As shown in Fig. 1, each node consists of an optical cross-connect (OXC) controlled by an IP router. Information about the availability of network resources is flooded in the network and distributed to all routers using the link-state protocol, *open shortest-path first* (OSPF). Every IP router is assumed to have a complete (not necessarily accurate) view of the network state.

A connection demand is referred to as a *session*. A Session $\psi_i = (s_i, d_i, B_i)$ is a triple where s_i is the source IP router, d_i is the destination IP router, and B_i is the bit-rate of the connection. A successful establishment of a session does not only entail logical aspects such as finding enough wavelengths in the network for the lightpath, but also considering the non-ideal behavior of the optical signal. As it will be discussed in the next subsection, an upper bound exists on the maximum length

0-7803-7206-9/01/$17.00 © 2001 IEEE

of an all-optical path and this upper bound is dependent on the bandwidth of the connection as well as the route taken.

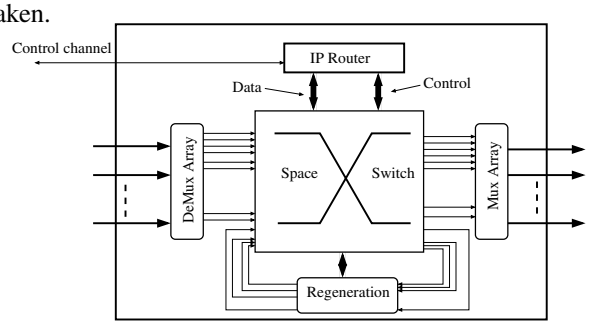

Figure 1: Node architecture. The node consists of an optical cross-connect (OXC) controlled by an IP router and capable of on-demand regeneration.

B. Transmission Impairments

Transmission impairments factors can be divided into two categories: *linear* and *nonlinear effects*. Linear effects are independent of the actual signal power and do not provide crosstalk between different wavelengths carried on the same fiber. Nonlinear effects are more difficult to handle and generate crosstalk between different channels. As an example, Polarization Mode Dispersion (PMD) (a linear effect) is one of the most important factors which limit the length of the transparent segment of the optical lightpath. When PMD is considered, a constraint is imposed on the time-averaged differential time delay between two orthogonal states of polarizations, $\Delta\tau$ [2]. This constraint requires that $\Delta\tau$ be less than a fraction of the bit-rate [4]. PMD can be thought of as providing an upper bound on the number of cascade amplifiers the optical signal can pass through without the need for regeneration. Equation 1 (adapted from [4]) provides the constraint on the length of the transparent optical path, where $path$ is a set of optical fibers, $Bandwidth$ is the bandwidth of the channel in $bits/second$, $D_{PMD}(l)$ is the fiber PMD parameter of Link l measured in $ps/\sqrt{km}$, and a has typical value of 0.1. Old fibers usually have PMD value equal to $0.5ps/\sqrt{km}$. Newer and improved fibers have PMD parameter equal to $0.1ps/\sqrt{km}$. It is expected that a given physical network will have a mixed set of links with different PMD parameters. Thus, it is wise to include this parameter in the link advertisement messages of OSPF.

$$Bandwidth \times \sqrt{\sum_{l \in Path} D_{PMD}(l)^2 \times Length(l)} \leq a \quad (1)$$

The constraint in Eqn. (1) will be used to illustrate the validity of the management approach.

II. Protocol Description

Figure 2 shows the generic format of the signaling message exchanged by the IP routers. This message is exchanged between neighboring OXCs and forwarded when necessary. The source and destination fields in the IP header are set to the sending OXC and receiving (neighbor) OXC, respectively. The protocol ID in the IP header is set to the new (unused) value for this new protocol. The 8-bit *Type* field indicates the type of this message. A message can be either a *connection setup*, an *accept*, a *reject*, a *release*, or a *tear down*. The 16-bit *Length* field is used to store the length (in bytes) of the signaling message excluding the IP header. The 16-bit *checksum* field covers the message excluding the IP header. A variable-sized data is exchanged between neighboring OXCs. Each IP router is assumed to

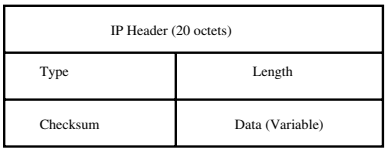

Figure 2: Generic format of the signaling message.

have knowledge of the topology including wavelength availability and transmission characteristics of optical fiber links. This information is distributed through the link-state protocol (e.g., OSPF). An IP router which sees the need to establish a high-capacity optical pipe to another IP router, calculates the constrained shortest path in the optical network. This path (along with wavelength and regeneration requests) are stored in a new control packet whose generic format is shown in Fig. 3.

As depicted in Fig. 3, the *sequence* field is a 32-bit number which is basically the connection ID and is unique within an OXC. The path which this message should go over is explicitly embedded in the message as a series of IP addresses starting with the source OXC IP address and ending with the destination OXC IP address. The number of OXCs on the path (including source and destination) is stored in the 16-bit *hop count* field. The *top* field is a 16-bit number pointing to the next hop. The *wavelength vector* is a series of wavelength numbers which this path is composed of starting from the source OXC. *Regeneration vector* is a bit vector indicating the request for regeneration at intermediate OXC. The need for regeneration can stem from either transmission impairments (to clean the signal) or to allow for opaque wavelength conversion.

After the source node computes the path and generates the wavelength assignment and regeneration request, It makes *top* point to *OXC1* and sends the IP packet to the first hop (OXC1). Each OXC receiving a connection

setup request does the following: It first increments the *top* field. If *top* points to the Destination OXC and it has the same IP as this OXC, then the message is forwarded to the control process. On the other hand, if *top* points to the Destination OXC and it is different from the current OXC, the message is silently discarded. Otherwise, the control message is read and the output wavelength specified by the *wavelength vector* is reserved for this connection. Any regeneration request is also satisfied using the local regenerators. After that, the control message is sent to the OXC whose IP is pointed to by *top*. This process of relaying the control message and reserving resources is repeated until the message is received by the destination of the connection OXC. When the destination OXC receives the connection request, it can either accept or reject the connection. If the connection is accepted, the destination prepares an *accept* message consisting of the reverse path. This message is then sent to the neighboring OXC (i.e, the last intermediate OXC). An OXC receiving an *accept* connection first looks to see if the destination field of the message is identical to its own IP. If it is not, the message is relayed to the next OXC. If the OXC IP matches the destination IP, the OXC starts transmitting on the configured path. A *reject* message is treated in the same manner except that the source OXC might attempt to reestablish the connection or report failure of the connection setup.

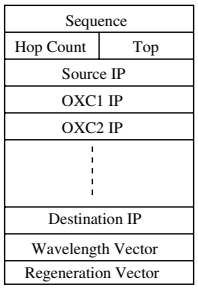

Figure 3: Format of the control packet.

Although the source OXC computed the path and wavelength assignment based on the current network status, the feasibility of the path might change due to network dynamics. If at an intermediate OXC the connection cannot be satisfied, a *release* control message is prepared which traverses the OXCs from this OXC and back to the source OXC. Every OXC receiving a *release* message, checks its own reservation table. If the connection (uniquely identified within the network by the triple $<$ Source, Destination, Sequence $>$) has been reserved, all resources including wavelengths and regeneration reserved for this connection are released and the release message is relayed to the next hop. Otherwise, the release message is silently discarded.

A. Link-State Distribution

The *open shortest-path first* (OSPF) protocol is extended for WDM-support. OSPF is a link-state protocol responsible for reliable flooding of link information throughout the network in the form of link state advertisements (LSAs). These LSA messages are delivered to all IP routers in the network; thus giving each IP router (after some convergence time) a complete view of the network topology and the current available resources. This information is valuable in the sense that source-based routing is used in dynamically provisioning optical channels.

B. Constrained Routing

Figure 11 shows an algorithm for creating a lightpath in the network. Given the current status of fiber links collected via the link-state protocol, the source node computes locally a path which traverses fiber links such that both the logical aspects (such as wavelength continuity constraint) and physical aspects (e.g., PMD factor) are not violated. It starts by creating a sub-graph consisting of all the links in the physical network having at least one wavelength available. After that, depending on the bit-rate of the connection, some links might not be able to handle the connection's bit-rate and these are also removed. A test is then made to check if the source and destination are in the same connected component. If they are disconnected, the call is rejected. Otherwise, Dijkstra's shortest path algorithm is run to find the shortest path with regard of the total PMD accumulation.

Next, the wavelength assignment process is executed. We utilize a variation of the First-Fit wavelength assignment scheme which runs as follows. The first available wavelength on the first fiber link is picked. Next, a wavelength is chosen from the next immediate fiber link with preference to the same wavelength as in the preceding link. If the wavelength is not available, the first wavelength available is picked and a regenerator (acting as an opaque wavelength converter at the connecting OXC) is installed. This process is repeated for all links in that order.

After taking care of the logical aspects of the routing and wavelength assignment problem, the physical aspects are handled as follows. The path is partitioned into the smallest number of transparent segments. These segments are then interconnected via regenerators.

III. Simulation Results

We extended the packet-based NS-2 simulator to support WDM wavelength-routed networks. The NS-2 simulator was chosen because of its extendibility features. The traffic model used is specified by three parameters:

ON, *OFF*, and *Period*, all measured in seconds. Period specifies the simulation time. Circuits for the node pair (s, d) are generated during Period as follows. At most one circuit is active at a time. The duration of each circuit is randomly and uniformly chosen from the ON parameter. The time between successive connection requests is randomly and uniformly chosen from the OFF parameter.

A. Convergence of Link-State Information

Clearly, the information available at each node and its accuracy have an effect on the performance of the network. In order to asses the convergence time of link-state information, consider the Italian network in Fig. 10 where every link is capable of carrying 8 wavelengths. Let us consider the following scenario. At Time $t = 30s$ of the simulation, every node in the Italian network requests 6 connections to each and every neighboring node. It is obvious that at Time $t = 30s$, the number of free wavelengths in the network is $72 \times 2 = 144$ wavelengths, since the number of unidirectional links in the network is 72. This information, however, has to be advertised to all nodes using the OSPF protocol. Figure 4 shows the convergence time of the link-state information starting from the time when there are no active connections in the network (i.e., available number of wavelengths in the network equals to $72 \times 8 = 576$). The figure shows two scenarios. The first scenario uses the original Italian network and the second scenario scales-up each link by a factor of 10. Two observation points are used: the first is at *Bolzano* node and the second at *Roma* node. Bolzano is picked as an observation point because it is a representative of an exterior node and Roma is used as a representative of a middle point in the network. For the original Italian network topology, we can observe that the time needed for the two nodes to receive the actual number of wavelengths in the network is very small (approximately 5ms). Notice also that the Roma node converges faster than Bolzano because of its position in the network. For the scaled-up case, the convergence time is increased to about $50ms$, which is what one can expect since the delay on every link is increased ten folds. However, the convergence of Bolzano is much faster than that of Roma in early stages of the link-state advertisements. This can be explained by looking at the Italian network topology and observing that the majority of the links are geographically closer to Bolzano than to Roma.

B. Regeneration Requirements

Consider the Italian network where every link has PMD parameter equal to $0.5ps/\sqrt{km}$. Traffic for the Italian network is randomly generated during a period of 1000 seconds and having both the ON and OFF parameters equal to 200 seconds. A heavy traffic with a total of 1784 connections is randomly generated. The same traffic is applied to the Italian network for different number of wavelengths. In the case of 8 and 16 wavelength on each link, 557 and 35 connections failed, respectively. All connections are successfully routed for the two cases of 24 and 32 wavelengths.

Figures 5, 6, 7, and 8 demonstrate the regeneration cost as a function of connection bandwidth for the four cases of wavelength count. In each figure, we show two scenarios for each bandwidth requirement. The first scenario (indicated by white bars) considers when all the nodes in the network are equipped with wavelength converters; making regenerators only used to clean up the signal. The second scenario (indicated by black bars) assumes no conversion in the network and wavelength conversion can be accomplished using regeneration. We make the following observations:

1. We notice that the PMD effect is only apparent for connection bandwidth of 5Gb/s or more.
2. We observe that as the number of wavelengths is increased, the need for regeneration for 5Gb/s connections is decreased until it reaches zero. This reduction is due to the fact that the number of good routes (in terms of PMD accumulation) is increased as we increase the number of wavelengths on every link; thus the constrained routing algorithm is able to find routes with regeneration cost equal to zero.
3. The average number of regenerators per connection for 10Gb/s traffic is slightly dependent on the number of wavelengths. The average number of regenerators per connection is 0.35 for 8 wavelengths and reaches 0.30 for 32 wavelengths.

C. Cost Reduction Through Smart Routing

To measure the cost reduction of the network (in terms of the electronic equipment) achieved by utilizing the PMD parameter in the route selection process, we conduct the following experiment. The same 1784 traffic used in the previous experiment and having 10Gb/s bandwidth is applied to the Italian network. All links in the network have 32 wavelengths allowing zero blocking.

We study the impact of PMD effect on the management of wavelength-routed networks. The most important aspect of this environment that needs to be studied is the impact of route selection on the overall cost of the network. To quantify this impact, we consider two routing schemes. The first scheme does not take into account the PMD parameter in the route selection and is

called *uninformed*. The other scheme does utilize this information in the route selection and is referred to as *informed*. The *informed* algorithm simply multiplies the PMD parameter of each link with its distance. Dijkstra's shortest path algorithm is run using this new metric.

Figure 9 shows the number of regenerators obtained using the two route selection schemes. On the x-axis, we show the percentage of unidirectional fibers having PMD equal to $0.1ps/\sqrt{km}$ (i.e., "good" fibers). On the y-axis, we show the total number of regenerators required for the given traffic. We first observe that the number of regenerators required when the *informed* route selection scheme is used is always less than or equal to the *uninformed* scheme. As expected, we notice that the two algorithms produce similar results in networks where one type of fiber is dominant over the other. On the other hand, the superiority of the *informed* scheme becomes more apparent in environments where the "good" fiber type is sparsely allocated in the network. This observation describes a realistic scenario given the fact that legacy networks are composed of the $0.5ps/\sqrt{km}$ PMD fibers and recent upgrades to regions of the network are conducted using new and improved fibers (i.e., $0.1ps/\sqrt{km}$ PMD parameter).

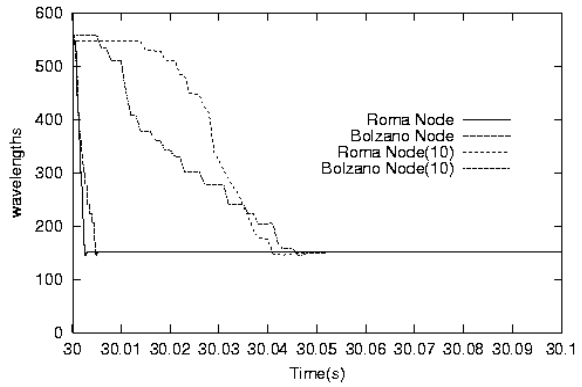

Figure 4: Convergence of the link-state information on the Italian network.

IV. Conclusions

In this paper we provided a detailed description of how the transmission impairments of the optical network can be accommodated for the dynamic provisioning of lightpaths. We show how one of the most important linear phenomena of optical transmission, Polarization-Mode Dispersion (PMD), is handled in the routing and management of optical networks. A constraint-based routing algorithm and a label-distribution protocol were developed. Simulation results on the Italian network topology were provided to assess the regeneration requirements.

We evaluated the impact of the transmission impairments on both the routing protocol as well as the overall cost of the network. Simulation results revealed that: 1) the feasibility of the lightpath is greatly affected by the PMD factor, especially for bit-rates exceeding 5Gb/s, and 2) significant improvement (in terms of cost) can be achieved through intelligent route selection schemes which take into consideration physical characteristics of fiber links. These cost savings are more apparent in networks with non-homogeneous fiber quality.

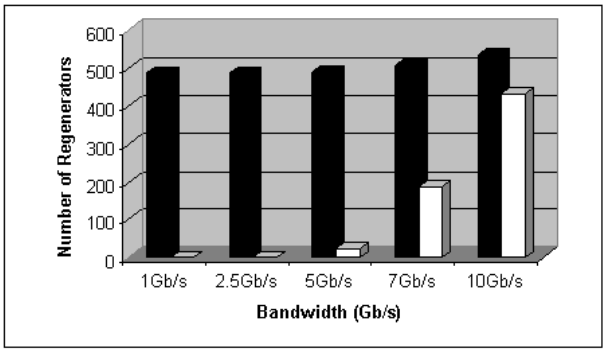

Figure 5: Total regeneration as a function of connection bandwidth. Eight wavelengths are used. Dark bars denote the case where no conversion is employed. White bars denote the case where wavelength converters are available on all nodes.

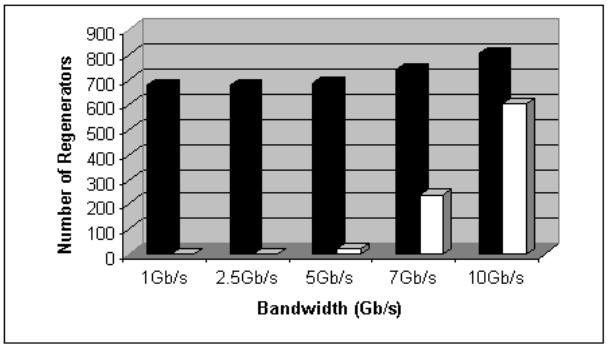

Figure 6: Total regeneration as a function of connection bandwidth. Sixteen wavelengths are used. Dark bars denote the case where no conversion is employed. White bars denote the case where wavelength converters are available on all nodes.

V. References

[1] M. Ali, *Transmission-Efficient Design and Management of Wavelength-Routed Optical Networks*. Kluwer Academic Publishers, 2001.

[2] R. Ramaswami and K. N. Sivaraja, *Optical Networks: A Practical Perspective*. San Francisco, CA: Morgan Kaufmann Publishers, 1998.

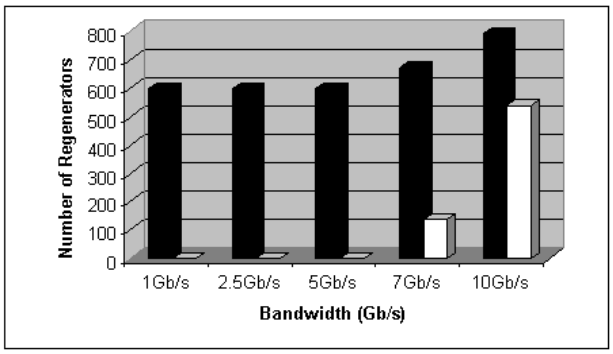

Figure 7: Total regeneration as a function of connection bandwidth. Twenty four wavelengths are used. Dark bars denote the case where no conversion is employed. White bars denote the case where wavelength converters are available on all nodes.

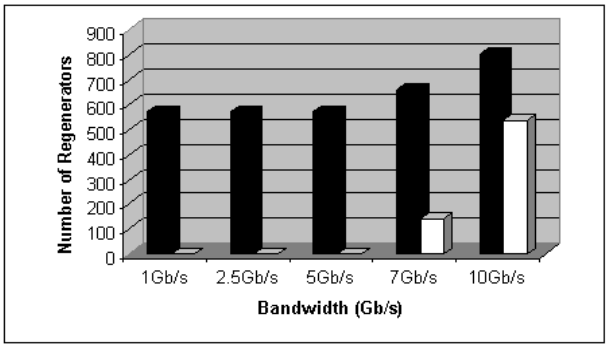

Figure 8: Total regeneration as a function of connection bandwidth. Thirty two wavelengths are used. Dark bars denote the case where no conversion is employed. White bars denote the case where wavelength converters are available on all nodes.

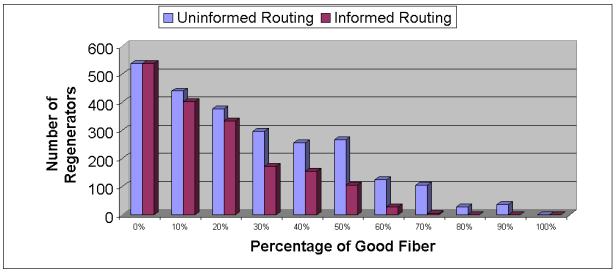

Figure 9: Number of regenerators required as a function of the percentage of "good" fiber. Graph shows the performance of the two route-selection algorithms.

[3] B. Davie and Y. Rekhter, *MPLS: Technology and Applications*. Morgan Kaufmann, 2000.

[4] J. Strand, A. L. Chiu, and R. Tkach, "Issues for routing in the optical layer," *IEEE Communications Magazine*, vol. 39, pp. 81–87, February 2001.

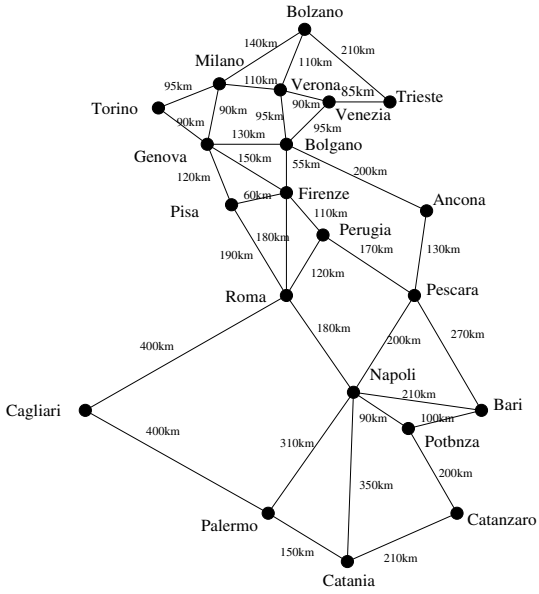

Figure 10: Topology of the Italian network.

Procedure *PrepareSetupMessage*(ψ_i)
BEGIN
Let *Network* be a subgraph of the original topology where an edge exists between two nodes iff at least one wavelength is free on the fiber link between them.
Remove from *Network* all edges which cannot handle bit-rate of Session ψ_i without regeneration.
IF s_i and d_i are disconnected in *Network* **THEN**
"Connection is blocked"
RETURN
END IF
Let *Path* equal to the shortest path (in terms of PMD accumulation) from s_i to d_i in *Network*
Use the First-Fit algorithm to assign wavelength.
Install regenerators if required.
Partition the path into segments where each segment has the maximum length without need for regen.
Install regenerators between segments.
Let *c_message* be a control packet where the three vectors: constrained path, wavelength assignment, and regenerator request have been filled.
RETURN *c_message*
END

Figure 11: Constrained routing algorithm with transmission impairments considerations.