J Netw Syst Manage (2008) 16:325–349
DOI 10.1007/s10922-007-9074-9

Policy-Based Grooming in Optical Networks

Fábio Luciano Verdi · Cláudio Carvalho · Maurício F. Magalhães · Edmundo R. M. Madeira

Published online: 18 August 2007

Abstract This work presents a discussion about policies and architecture to aggregate Internet Protocol/Multiprotocol Label Switching (IP/MPLS) traffics within lightpaths. The scenario is that of IP/MPLS client networks over an optical network. It is well known that aggregating lower traffic flows (e.g., packet-based LSPs—Label Switched Path) within higher traffic flows (e.g., lambda-based LSPs) is considered an effective way to maximize the use of the optical network resources. In this work, the policies are divided into two groups. The first one, which solely considers the class of the flow (High Priority—HP or Low Priority—LP), consists of simple policies meant to aggregate packet-based LSPs within lightpaths. In this group, the policies we have defined intend to reduce the optical network overhead to remove and reroute LP LSPs. The second group presents more sophisticated policies taking into account the possibility of having to deal with further transport faults. In this case, the grooming is better planned and the defined policies tend to reduce the negative impact when a failure is detected in the optical transport network. Our approach has been implemented to validate the policies and the results for each group are showed and discussed.

Keywords Traffic grooming · Policy-based network management and optical network management

F. L. Verdi (✉) · M. F. Magalhães
DCA-FEEC-UNICAMP, Campinas 13083-970, Brazil
e-mail: verdi@dca.fee.unicamp.br

M. F. Magalhães
e-mail: mauricio@dca.fee.unicamp.br

C. Carvalho · E. R. M. Madeira
Institute of Computing, (IC-UNICAMP), Campinas 13084-971, Brazil

1 Introduction

A typical optical network ranges from ten to thousands of Gb/s of available bandwidth. These networks are likely to be consisted of elements such as routers, switches, Dense Wavelength Division Multiplexing (DWDM) systems, Add-Drop Multiplexers (ADMs), Photonic Cross-Connects (PXCs) and Optical Cross-Connects (OXCs) [1]. Due to the advent of Generalized Multi-protocol Label Switching (GMPLS), there is a common sense to use it to dynamically provide resources and perform the routing and signaling functions of the control plane.

The idea of label switching in Multiprotocol Label Switching (MPLS) only considers the packet domain taking into account Internet Protocol (IP) networks. However, due to the recent growth in optical technologies and their high bandwidth, many packet-based network flows are expected to be nested within lightpaths to cross the optical domain and reach their destination. Such lower order Label Switched Paths (LSPs) aggregation within higher order LSPs is well known as the traffic grooming problem. Lightpaths are seen either as LSPs or optical LSPs (hereafter optical LSP and lightpath will be used interchangeably) and because of technologies like DWDM it is now possible to have a very large number of parallel links between two adjacent nodes (hundreds or even thousands of wavelengths if multiple fibers are used). Thus, GMPLS has emerged as a solution to act not only in the packet or cell based networks but also in the time division networks (using Time Division Multiplexing—TDM), where the switching decision is based on time slots, and optical networks, where the switching decision is based on wavelengths or physical ports.

Although GMPLS considers all the data forwarding kinds above, IP over DWDM is the one that has emerged recently. In this context the overlay model is very much recommended for service providers (e.g., Telecom companies) considering their great interest in acting as transport networks for IP client networks. A very typical and promising scenario is to have MPLS client networks with their packet-based LSPs asking for an optical resource (typically an optical LSP) so that they may cross the optical domain and get to their destination.

Depending on how the aggregation of packet-based flows within lightpaths is done, the use of the network bandwidth may be either maximized or wasted. It is clear that if some rules are followed the optimization of the network resources may be boosted allowing thus more traffic. Besides, considering that packet-based flows can be divided into different classes (e.g., DiffServ), it is necessary a thorough treatment for higher classes. In this paper we present some simple policies to efficiently improve the use of the optical network bandwidth, the architecture to apply such policies and the results of the simulations we have performed.

This work proposes a set of policies to manage the installation and aggregation of packet-based LSPs within optical LSPs assuming that there are several lightpaths between two end nodes. The policies are divided into two groups. In the first group we defined simpler policies intending to aggregate IP/MPLS traffic considering only the class of the flow. We worked with two classes of traffic, Low Priority (LP) LSPs and High Priority (HP) LSPs. The HP LSPs classes have higher priority than LP LSPs so that if an HP LSP needs to be installed in an optical LSP, some LP LSPs

from that optical LSP will have to be removed to attend the higher requirement in the case the lightpath does not retain enough bandwidth. These removals are very expensive on optical networks due to the overhead necessary to tear down one or more LP LSPs and reroute them in another optical LSP. On account of that overhead, reducing the number of removals is a very critical and important point to be considered in optical networks. Two previous works were done considering the policies for the first group [2, 3]. The second group of policies is more complex and sophisticated. We extended the previous policies and created new ones which take into account the aggregation of flows within a lightpath to reduce the impact of a failure. The policies try to aggregate the IP/MPLS traffic in such a way that when a given failure happens, the number of affected packet-based LSPs is smaller when compared to a non-policies scenario. Another previous work [4] considering the policies for failure has been done using a simple network topology. In this current work however, we use the National Science Foundation Network (NSFNet) topology to reflect a more real scenario. It is worth to remark that the second group of policies does not exclude the first one but is rather an extension of it, as mentioned earlier.

The aggregation of traffic is dynamically done by a Policy Manager (PM). For each request[1] that arrives, the PM looks for a lightpath that can accommodate the flow. If a lightpath is found assuming all the constraints specified by the flow, that flow is then groomed in the lightpath, otherwise the request is refused.

The architecture proposed in this work consists of an Admission Control (AC) responsible for receiving a request and verifying the Service Level Agreement (SLA), a PM responsible for applying the policies and finding an optical LSP to accommodate the packet-based LSPs, a Fault Manager (FM) for managing failure events and a Resource Manager (RM) for controlling the resources. Each module will be described in detail further in this work.

This paper is organized as follows. In the next section some related works are described. Section 3 shortly explores some basic concepts and the reference scenario. Section 4 presents our proposed architecture and its modules. Section 5 is dedicated to depict the defined policies and the results of our simulations. Finally, Sect. 6 presents the conclusion.

2 Related Works

Most of the recent works [5, 6] has dealt with heuristics for off-line routing, multilayer networks and multihop traffic grooming. They have not considered the use of policies to accommodate new on-line (dynamic) requests in the overlay model taking into account QoS and the overhead to remove and reroute packet-based LSPs.

In [5], a traffic engineering system is presented considering the multilayer approach and taking into account both methods of routing, off-line and on-line. In

[1] In this work, the term request represents an IP/MPLS flow and its QoS requirements. They will be used interchangeably.

[7], the traffic grooming problem is well treated and a formulation on how to use an integer linear programming is presented.

The traffic grooming and network survivability are issues that have been discussed for optical WDM networks. Many works use the traffic grooming technique to optimize the use of network resources and, consequently, reduce the cost with network equipments. Zhu [8] investigates the problem of efficiently provisioning connections of different bandwidth granularities in heterogeneous WDM mesh network through dynamic traffic grooming schemes under traffic engineering principles. Different traffic engineering issues are presented. Yao [9] studies the traffic grooming problem in WDM mesh networks using the fixed alternate routing. An algorithm to support on-line provisioning of multi-granularity subwavelength connections is proposed, the Fixed-Order Grooming (FOG). Two groups of grooming policies are also proposed. The first group is formed by three policies responsible for addressing the route selection problem. The second group has three policies as well and addresses the wavelength and transceiver constraints.

Other works on traffic grooming are also focused on network survivability. Canhui [10] proposed two methods for traffic grooming: the Protection-at-Lightpath (PAL) and the Protection-at-Connection (PAC). These methods are different in terms of routing and the amount of resource required and similar in terms of provisioning an end-to-end protection. Considering the occurrence of a single failure, in PAL the end nodes of the lightpaths switch the traffic to the respective backup lightpaths and in PAC the end nodes of the connections switch the traffic to the respective backup lightpaths. The authors try to optimize the usage of network resources through the traffic grooming.

The survivability became a critical concern in optical WDM networks. Many works address the problem through recovery mechanisms. Ramamurthy [11] examines different approaches to protect WDM optical networks from network element failures. Based on the protection and restoration paradigm the study investigates the wavelength capacity requirements, and routing and wavelength assignment of primary and backup paths. The protection-switching time and the restoration time are also investigated. The results showed that there is a trade-off between the capacity utilization and susceptibility to multiple failures.

Alanqar [12] presented an overview of the fault management mechanisms involved in deploying a survivable optical mesh network and described different parameters that can measure the quality of service provided by a WDM mesh network to network clients (e.g., ATM network backbones and IP network backbones). Some of those parameters are service availability, service reliability, restoration time and service restorability. Considerations on how to improve these parameters are also discussed.

Fawaz [13] proposed a SLA applied for optical networks (O-SLA). The authors describe three classes of services and some traffic and performance parameters to compose each class, for example, the connection setup time, service availability, resilience and routing constraints. The use of O-SLA for applying policies in optical networks is also discussed.

Although some works deal with grooming, multilayer integration and survivability, to the best of our knowledge none of them addresses the failure problem

during the traffic admission. Furthermore, the works cited above deal with traffic grooming and/or survivability issues. However, none of them addresses the issues altogether using the Policy Based Network Management (PBNM) paradigm. The PBNM is becoming a promising paradigm to efficiently allocate network resources for client networks. Some motivations for this paradigm can be found in network scenarios where the SLAs contain various client requirements and the network resources can be configured to offer different levels of quality of service.

3 Network Scenario and Technical Background

The concept of nesting LSPs (LSP within LSP) has been already used in MPLS networks. In those type of networks, lower bandwidth LSPs having the same destination node are aggregated within another higher bandwidth LSP by means of label stacking. Considering that in optical networks the bandwidth of an optical LSP is very high (from one to hundreds of Gb/s), lower bandwidth LSPs can be nested within the lightpath.

The concept of elasticity is also very important since it allows a packet-based LSP to increase or decrease its bandwidth as needed. Because transport networks will charge the clients for the used bandwidth, having a mechanism to give bandwidth to the connection only for the time it is actually requested is very important. As the bandwidth is reduced, the portion that is released is left at the disposal of the network. The possibility of increasing and decreasing bandwidth raises a problem when classes of priority are defined. The problem occurs when a higher priority packed-based LSP wishes to increase its bandwidth. If there is available bandwidth in the lightpath, no LP LSPs will need to be removed. However, when no bandwidth is available, one or more LP LSPs will have to be removed to accommodate that new demand. This is a typical problem that comes across whenever grooming is done in optical networks. Every slice of bandwidth that is idle within a lightpath should be filled to increase the provider revenue. The way these idle spaces are filled drives the usage of the optical resources. That is similar to the memory fragmentation problem in which the sum of all idle memory spaces is high, however no single space is capable of accommodating data. If the grooming is done without some ruling, the bandwidth of each lightpath will not be used efficiently and a small aggregation of traffic may occur causing connections to be refused. When the traffic is accommodated using some policies, the waste of resources is reduced. Moreover, when failures are considered, the policies should have a view of the physical topology to attend protection constraints required by each flow.

Figure 1 shows an optical network and its IP/MPLS network clients. There are three optical LSPs between the two border nodes. These LSPs represent a logical topology and are considered as Forwarding Adjacencies (FA) since each one is advertised as a Traffic Engineering (TE) link into Intermediate System–Intermediate System (IS–IS) or Open Shortest Path First (OSPF) protocols allowing other Label Switched Routers (LSRs) to use the FAs for their path computation [1]. This is a clear way to maximize the use of the optical network resources.

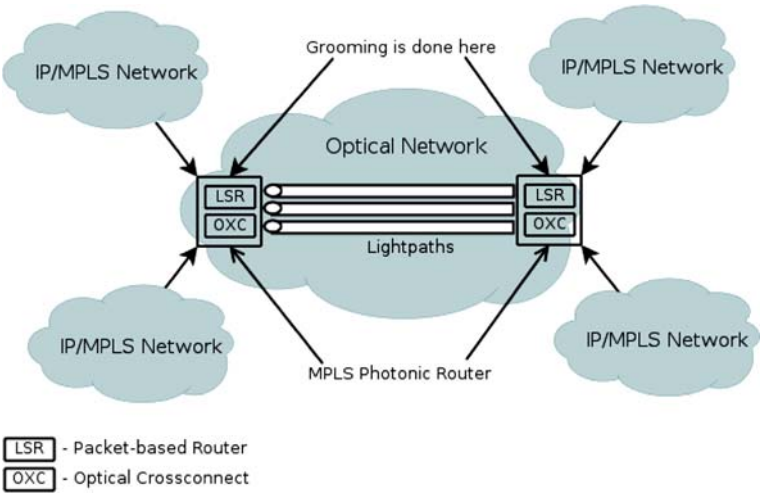

Fig. 1 The overlay reference scenario

Figure 1 also illustrates a typical scenario with MPLS networks acting as client networks and the optical network acting as the transport network (IP over optical network integration). This integration between IP/MPLS and optical layers is done by means of a standard User-Network Interface (UNI) [14]. In this case, the IP/MPLS asks for a connection and the optical layer tries to set it up according to the SLA. The three optical LSPs shown in Fig. 1 are seen by client networks as tunnels from the ingress node to the egress node and, therefore, from the point of view of the client networks, these nodes are peers. If no policy is applied, the packet-based LSPs will be aggregated taking into account only the available bandwidth in each tunnel.

Although the optical network solves many known problems, it brings new challenges for the research community. One of the main problems deeply analyzed is finding how to minimize the impact of failures in the network. Since each link has a high bandwidth, a failure in a link will cause a lot of data loss. Much effort has been done on trying to use the same idea of SONET/SDH networks whose recovering time is about 50 ms. However, it is very difficult to reach such time in a meshed optical network. The IETF has defined the GMPLS architecture by extending some protocols already used in MPLS networks. Those protocols have been defined for dealing with failures treatment. An example of that is the *Notify* message defined in the Resource Reservation Protocol (RSVP) that was extended to support GMPLS networks [15]. There also have been some attempts related to inter-domain protection [16] but nothing has been defined as standard so far.

The research community has defined (not formally) four main types of protection. The most basic and simplest one is the self-explained unprotected traffic. On the other extreme is the 1 + 1 protection which defines that for each primary lightpath there is exactly one dedicated backup lightpath carrying the same

traffic at the same time. The egress node selects the best signal to be dropped.[2] In case of a failure, only the egress node needs to switchover to the backup. In between these two levels, there are two levels of protection named 1:1 and 1:N. In the 1:1 scheme, the traffic is only sent in the primary lightpath and the backup lightpath can be used for extra traffic. When a failure affects the primary lightpath, the extra traffic being transported on the backup needs to be blocked and the traffic from the primary lightpath is preempted to the backup. The switchover is performed in the ingress node as well as in the egress node. The last scheme of protection is the 1:N which defines that there is only one backup lightpath for N primary lightpaths. If one of these primary lightpaths comes to fail, the remaining N−1 primary lightpaths become unprotected until the failure is repaired. Further details about recovery can be found in [17].

In this work, the first group of policies was tested by using a virtual topology connecting two border nodes. This topology only shows the established lightpaths between these two nodes as shown in Fig. 1. This is enough to apply the first group of policies. However, the failure policies need to have a view of the network physical topology. To perform this simulation, the NSFNet network topology was adopted. This will be explained in details in Sect. 5.

4 The Proposed Architecture

The architecture proposed in this work is composed of four management modules: AC, FM, PM, and RM. There is also a Policy Repository responsible for storing the policies. These modules were designed in order to get a basic infrastructure to apply policies in optical networks as well as to control all the necessary information for the management of the IP/MPLS over DWDM integration [18]. The architecture is presented in Fig. 2 and a brief explanation of each module is presented right below.

- *Admission Control* (AC): The AC receives the requests sent by the IP/MPLS networks and prepares them, loading the lightpaths (from the RM) between the source/destination pair. After getting all the lightpaths that connect the ingress and the egress nodes, the AC sends such information to the PM which in turn is responsible for applying the policies (see below). The AC module is also in charge of resending to the PM the traffic flows that were blocked during the admission phase in a tentative of readmitting them;
- *Policy Manager* (PM): The PM implements the policies by analyzing a pool of candidate lightpaths (received from the AC), trying to find one with available resources to accommodate a given IP/MPLS request. Also, the PM is responsible for receiving a pool of failed lightpaths from the FM in order to try to readmit them by following specific failures policies;
- *Fault Manager* (FM): The main function of the FM is to receive the link failure events generated by the optical network equipments and prepare the lightpaths

[2] In optical networks the term *drop* means to extract the optical signal from the network and make it ready for further processing.

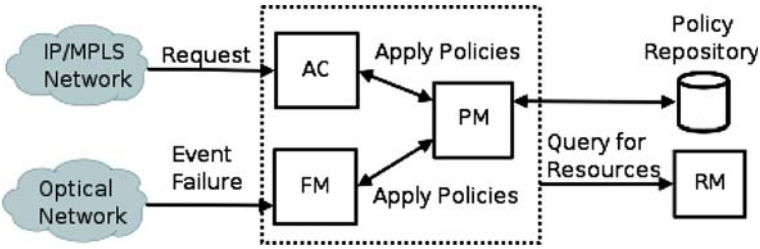

Fig. 2 The proposed architecture

contained in the failed fiber[3] by separating them into groups of lightpaths according to their type of protection. Then, the FM sends each group of lightpaths to the PM which in turn applies the specific defined policies for failures treatment;

- *Resource Manager* (RM): The RM is responsible for storing the information about the virtual and physical topologies. It is accessed by the AC, FM and PM in order for them to obtain any kind of data related to the resources and policies.

The policies were created to try to separate HP LSPs from LP LSPs. The main point is to dynamically find a lightpath to allocate each packet-based LSP in order to reduce the number of LP LSP removals in case of increasing the bandwidth of HP LSPs and reduce the negative impact when a failure is detected in the optical network. This decision is taken each time a new packet-based LSP (L_i) is being installed. As we can see in Fig. 2, a request arrives in the AC and carries the following data:

- *Request bandwidth*: This is the quantity of bandwidth required by the LSP at the moment of its installation. After it is installed, this value may be either increased (up to the maximum bandwidth) or decreased;
- *Maximum bandwidth*: This is the maximum bandwidth an LSP can ask for, based on what was agreed in the SLA;
- *Level of protection*: 1 + 1, 1:1, 1:N, unprotected (only used in the second group of policies).

We did not specify the format and the way a request comes from the client network since, in this work, we are assuming that the request comes from an IP/MPLS network and as such, it can be a PATH RSVP message that is received by the UNI. However, such request could come through a management interface or any other type of interface defined by a provider. The establishment of lightpaths in each domain can be done using GMPLS, or Automatically Switched Optical Network (ASON) [19] or any other solution. Our architecture does not depend on the type of technology used to set up the lightpaths. It should be highlighted that the way the lightpaths are established is up to each local administrative domain.

[3] In this work we are considering only fiber failures. Other types of failures such as wavelenghts failures and OXC failures are not considered.

Figure 3 shows the partial class diagram used to model the policies that will be described in the next section. This class diagram is based on the Common Information Model (CIM) specification [20], which defines an approach on how to apply policies using actions and conditions. In Fig. 3 the conditions as well as actions can be combined (compound) using the classes *CompoundPolicyCondition* and *CompoundPolicyAction* in order to attend specific policies. When removals are needed, the actions *PolicyActionRemoveLSP* and *PolicyActionInstallLSP* are combined to support the policies that triggered the LP LSP removals. Other compositions can be done based on each specific policy. Furthermore, policies can be grouped and organized in sets to represent specific rules.

The class diagram presented in Fig. 3 allows the PM to manage the resources of more than one domain. Figure 4 illustrates how such multi-domain management could be done. Firstly, the *PolicySet* is created to store the *PolicyGroups* of each domain. With the *PolicyGroup*, the policies (*PolicyRule*) of a given domain can be structured in accordance to the requirements of the network administrator. In a simpler case, this structure can be formed by only one level, such as in Fig. 4. In a more complex case, groups that contain *PolicyGroups* can be created resulting in

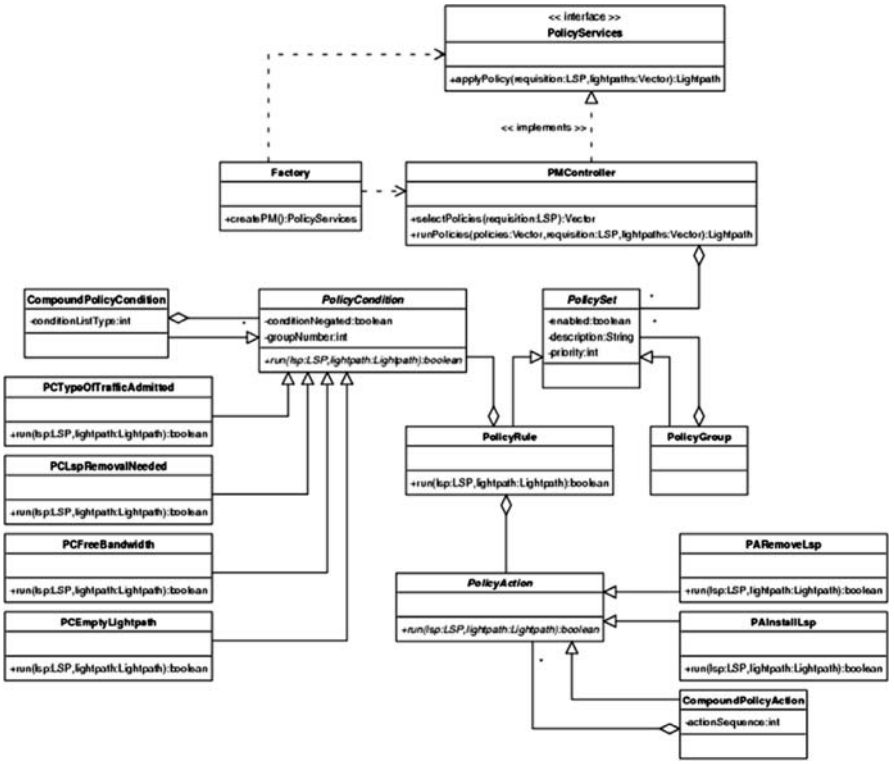

Fig. 3 Policy information model and its association with the system classes

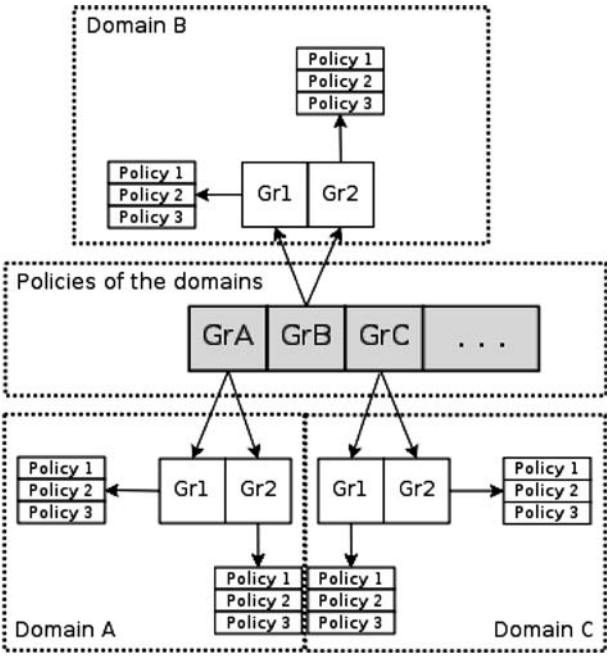

Fig. 4 Structure of policies

many levels of policies. In this paper, our method considers only one domain and one level of policies. Multi-domain management is left for further study.

Next section presents the proposed policies and the results obtained in our simulations.

5 Detailing the Policies and the Results

As mentioned earlier, the policies were divided into two groups. The first one only considers the class of the flow as being HP or LP. The second group takes into account the level of protection.

In order to better comprehend the graphs, we first show the transition flow that represents the state of an IP/MPLS flow (see Fig. 5).

The initial state represents the request arrival. From the initial state, the request can be admitted, refused or finished. The finished state refers to the state by which the flow terminates in a regular way (e.g., user ends the transmission). If the flow is admitted, it may be moved to the removed state that is an intermediary state whereby a new decision needs to be taken. From that state, the flow may be blocked (it could not be aggregated in another lightpath) or readmitted (the flow has been removed and could be aggregated in another lightpath). From the readmitted state the flow may be removed again and the loop either continues or the flow can be

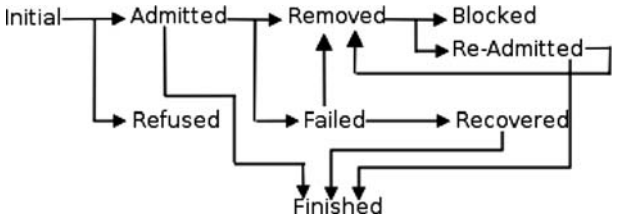

Fig. 5 The transition flow of an IP/MPLS request

finished. Back to the admitted state, the flow may fail (failed state). The failed state means that the flow is located within a lightpath whose fiber failed. Then it can be recovered which means that either it had been previously protected and after the failure it was directly switched over to its backup, or it had been removed (unprotected traffic) continuing the loop as before (from the removed state). From the recovered state, the flow can be finished.

5.1 Detailing the First Group of Policies—Without Polices for Failures

A given policy will be applied only if its precedent failed, e.g., policy 2 is applied only if policy 1 fails and so forth. The allocation of an HP LSP is done taking into account only the maximum bandwidth in order to allow the HP LSPs to ask for more bandwidth after the installation. Thus, in case of being an HP LSP, the verification of the available bandwidth in a given lightpath should consider the maximum bandwidth of all HP LSPs already installed in that lightpath (an example is showed when policy 2 is presented below). The request bandwidth of the HP LSPs is only considered when an LP LSP is being installed. While the HP LSPs do not ask for more bandwidth, the bandwidth that is left (the difference between the maximum bandwidth and the bandwidth that HP LSPs are actually using) can be used by LP LSPs. Therefore, when an LP LSP is to be allocated, not only the request bandwidth of all LP LSPs already installed needs to be verified but also the request bandwidth of the HP LSPs also needs to be considered since the sum of all of them represents the quantity of bandwidth that is actually being used at that moment in a given tunnel. Based on that sum, it is possible to know if there is enough bandwidth to accommodate the LP LSP. The proposed policies are presented below.

Let R be the Request and L be a given lightpath between an ingress node and an egress node.

if R is HP (the allocation is done based on the maximum bandwidth of all already installed HP LSPs):

- *Policy 1*: Accommodate R in L if L is not empty and has only HP LSPs. This policy is always the first one to be applied and it assumes that if a lightpath has only HP LSPs, when those HP LSPs require more bandwidth no LP LSP removals will be needed;

- *Policy 2*: Accommodate R in L if L is not empty and has both kinds of LSPs. This policy has two restrictions and both must be matched:

(1) Accommodate R in L if no LP LSP removals are needed at this moment, i.e., at the time the LSP is being installed;
(2) Accommodate R in L if the sum of the maximum bandwidth of all HP LSPs and the sum of the request bandwidth of all LP LSPs do not exceed the bandwidth of L. Thus, if in the worst case all HP LSPs in the tunnel need to increase their bandwidth, no LP LSPs will be removed. As an illustration, we may suppose that a new request to install the following HP LSP arrives: request bandwidth is 100 Mb/s and maximum bandwidth is 200 Mb/s, and suppose that the bandwidth of L is 1 Gb/s and has the following already installed LSPs:

    (a) HP LSP 1: Request bandwidth 100 Mb/s and Maximum bandwidth 200 Mb/s;
    (b) HP LSP 2: Request bandwidth 200 Mb/s and Maximum bandwidth 300 Mb/s;
    (c) LP LSP 3: Request bandwidth 100 Mb/s and Maximum bandwidth 200 Mb/s;

The sum of the maximum bandwidth of all HP LSPs (including the one that is being installed) is 700 Mb/s and the sum of the request bandwidth of all LP LSPs is 100 Mb/s. Since the total sum is 800 Mb/s and the lightpath has 1 Gb/s, the LSP can be installed in L. However if we suppose that the maximum bandwidth of the HP LSP 1 is 400 Mb/s and the maximum bandwidth of the LSP being installed is 300 Mb/s, the LSP would not be allocated in lightpath L since the total sum would be 1.1 Gb/s.

- *Policy 3*: Accommodate R in L if L is not empty and has both kinds of LSPs. This policy does not comprise the policy 2's restriction (b);
- *Policy 4*: Accommodate R in L if L is not empty and has both kinds of LSPs. This policy is the same as policy 2 however without restrictions (a) and (b);
- *Policy 5*: Accommodate R in L if L is empty.

if R is LP (the allocation is done based on the sum of the request bandwidth of all already installed LP LSPs and HP LSPs):

- *Policy 6*: Accommodate R in L if L is empty;
- *Policy 7*: Accommodate R in L if L is not empty and has only LP LSPs;
- *Policy 8*: Accommodate R in L if L is not empty and has both kinds of LSPs;

In case policies are applied and a lightpath is not found, the LSP will be blocked and will not be installed. Note that an LP LSP installation does not cause any removal since there is only one level of priority among the LSPs of this type. Moreover, an LP LSP is allowed to increase its bandwidth only if there is the quantity that it is requiring in the lightpath. When applying policy 4, there can be more than one available lightpath to allocate the LSP. In that case and considering that there will

be LP LSP removals in each lightpath, we choose the lightpath which presented the smallest number of removals, thus reducing the overhead to tear down and reroute such LP LSPs.

It is important to mention that there are some different ways to aggregate packet-based LSPs within lightpaths without using policies. Basically, for this particular group we have assumed that when policies are not considered, the aggregation is done by sequentially looking for a lightpath until the one that has enough bandwidth to accommodate a given LSP is found.

5.1.1 Results of the First Group

Different scenarios were created to test the policies, each one using a different sequence to apply them. The results showed in this section consider the sequence which achieved the best results, which is the same sequence presented in Sect. 5.1 above: from policy 1 to policy 5 for HP LSPs and from policy 6 to policy 8 for LP LSPs. We assume that the bandwidth of each lightpath is 1 Gb/s, the number of LSPs to be installed is 200 and that about 50% of LSPs will ask for increasing their bandwidth. The LSPs are randomly created; their minimum bandwidth being 50 Mb/s and their maximum bandwidth 400 Mb/s. As a consequence, the request bandwidth (see Sect. 4) varies from 50 Mb/s to 400 Mb/s. The number of available tunnels, and consequently the quantity of available bandwidth between two end nodes varies from 10 to 40 (10–40 Gb/s). The simulation was run 300 times and the average was obtained after those 300 loops.

In the first simulations more HP LSPs ($\sim$66%) than LP LSPs ($\sim$33%) were created. This case represents a situation in which there are more HP LSPs than LP LSPs. The request bandwidth average for HP LSPs is 16 Gb/s. The maximum bandwidth average for HP LSPs is 35 Gb/s and the request bandwidth average for LP LSPs is 8 Gb/s.

Figure 6 depicts the number of installed LSPs as the bandwidth increases. With 40 Gb/s all the LSPs are installed: $\sim$133 HP LSPs and $\sim$66 LP LSPs (out of 200). The numbers are the same for both scenarios, either applying or not the policies. Figure 6 intends to show that the quantity of installed LSPs using the policies is the same as without them. The difference takes place when the HP LSPs ask for increasing their bandwidth. In that case when the defined policies are applied the number of removals is quite smaller as can be seen in Fig. 7 (increasing the bandwidth in 50%).

While the number of removals without using policies increases up to 19 and keeps this behavior until the end, the number of removals when applying the policies decreases and with 30 Gb/s there is no more removals. When 20 Gb/s of bandwidth is available in the optical layer, the number of removals using policies is 6 and 19 when they are not applied, a difference of 68%. The point is that without the policies, the more LSPs are accepted the higher the number of removals. However, when policies are applied, that is not the case because they separate HPs from LPs. Figure 8 shows the same situation except that the bandwidth is now being increased to the maximum value allowed. Since more bandwidth is now being

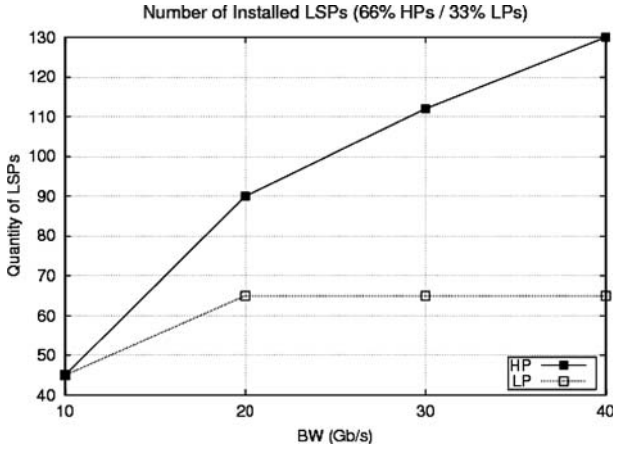

Fig. 6 Installed LSPs: ~66% HPs and ~33% LPs

required, the number of removals increases in both situations. With 20 Gb/s there are 35 removals using the policies and 42 without them, a difference of 16%. With 30 Gb/s there are 10 removals with policies and 52 without using them, a difference of 80%. Finally, with 40 Gb/s the number of removals is only one when applying the policies and 52 without using them.

The remaining graphs represent the scenario in which the number of HP LSPs and LP LSPs are about the same, 50% for each type. In this case, the request bandwidth for HP LSPs is 12 Gb/s, the maximum bandwidth for HP LSPs is 26 Gb/s and the request bandwidth for LP LSPs is 12 Gb/s. Figures 9 and 10 show respectively the results after increasing the bandwidth in 50% and to the maximum. We can observe that in Fig. 9 the number of LP LSP removals is greater when

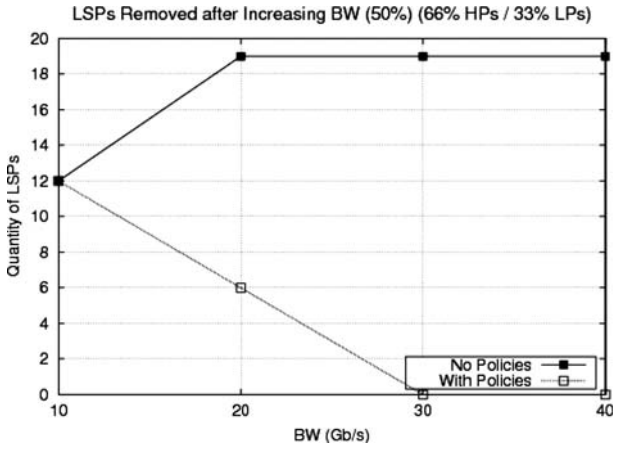

Fig. 7 LP LSPs removed after increasing the bandwidth in 50%

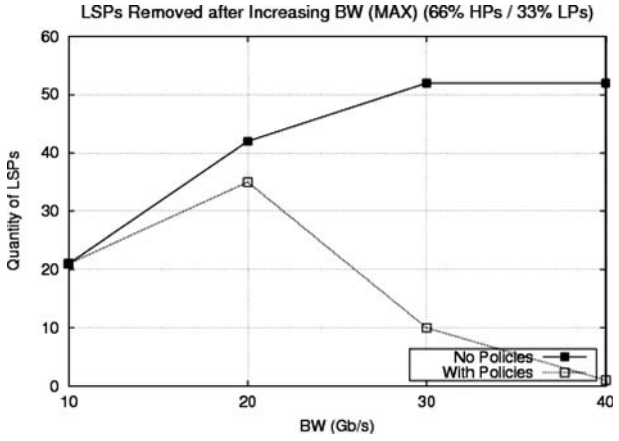

Fig. 8 LP LSPs removed after increasing the bandwidth to the maximum

compared to Fig. 7 because there are more LP LSPs in the network. With 30 Gb/s the quantity of removals is 25 without using policies and three with policies, a difference of 88%. The result is better with 40 Gb/s: no removals with policies and 25 without them.

In Fig. 10 we can see that because there is more demand for bandwidth (increasing to the maximum) more removals are necessary when compared to Fig. 9. That is the same case when Fig. 10 is compared to Fig. 8. Since there are more LP LSPs in the network, more LP LSP removals will be needed.

The sequence in which the policies are applied can generate different results. We verified that the sequence used in our simulations (shown in the beginning of this section) is the best sequence to be used. Another possible combination is the one in which policy 5 is applied before policy 2. By applying this sequence, the number of

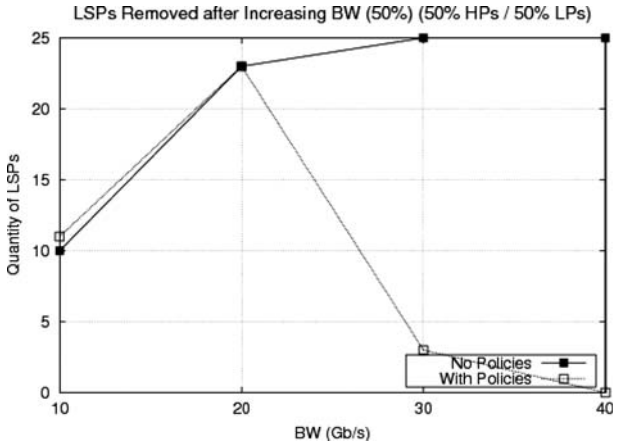

Fig. 9 LP LSPs removed after increasing the bandwidth in 50%

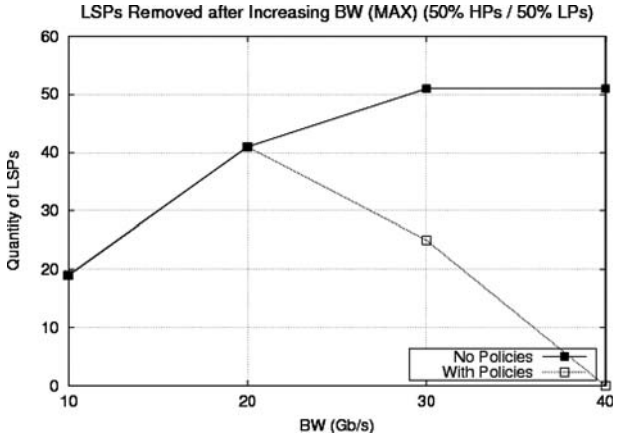

Fig. 10 LP LSPs removed after increasing the bandwidth to the maximum

removals after increasing the bandwidth increases up to 10%, but in several cases the quantity of removals is the same as the results presented in this section. Another possible sequence is applying policy 7 before policy 6. However this combination turned out to be not feasible since the number of removals during the installation of the LSPs increases in 60% in some cases.

The results presented above represent two scenarios. The first one is when there are 66% HPs and 33% LPs and in the second one there are 50% HPs and 50% LPs. We have tested more scenarios and observed that if there are many LP LSPs to be installed (more than 50% of LPs) the policies will be useful only when the network bandwidth is 20 Gb/s. On the other hand, if there are less than 30% of LP LSPs to be installed the policies start helping when the network bandwidth is 10 Gb/s. As a result, for those scenarios where traffic flows are mostly composed of HP flows (e.g., typically in multimedia scenarios) the use of policies is indicated.

5.2 Detailing the Second Group of Policies—Policies for Failures

For this group, we have developed three subgroups of policies, named G1, G2 and G3. It is important to highlight that when dealing with failures, the policies defined in the G1, G2 and G3 subgroups accommodate each IP/MPLS flow within a lightpath intending to reduce the impact of a given failure in the future. When the failure happens in the optical network transport there is not much to be done since the traffic was aggregated during the AC, therefore, after the failure takes place, the only feasible procedure is to preempt the protected flow and, as an extra attempt, try to readmit some failed traffics. Note that the tentative of readmitting traffic is done by resending the failed traffic to the PM and let it apply the policies. So, the ultimate effort of this second group of policies is done during the AC as for the policies of the first group, explained in the previous section (Sect. 5.1). Below, each subgroup of policy is explained separately.

- *Policy Group 1 (G1)*: This group is the simplest admission policy group. When a request arrives in the PM, it tries to install the request in a lightpath that offers exactly the same protection as required. It does not consider the class of service of the request. The G1 group is regarded equivalent as to not having policies since the policies defined in this subgroup are very simple (only the required protection is offered). Then, the G1 subgroup is our baseline parameter for comparison among the other subgroups.
- *Policy Group 2 (G2)*: It has an intermediate complexity. Its approach is to admit an LSP in a lightpath whose protection level matches the level of protection required by the request. Also, it always tries to keep together within lightpaths the LSPs which have the same class of service (HP and LP). This group of policies can be better explained as follows: Let R be the Request and L a given lightpath.
  - if R is Unprotected
    - if R is HP
      (a) Aggregate R in an unprotected L if the LSPs already aggregated in L have the same class of service of R;
      (b) Aggregate R in an unprotected L that is empty;
      (c) Aggregate R in an unprotected L. That L might have both LP and HP LSPs;
      (d) Aggregate R in an unprotected L if the removal of one or more LP LSPs of L releases enough bandwidth to install R;
    - if R is LP
      (a) Repeat the three first steps described above for HP;
      (b) Aggregate R in a backup L that is not empty;
      (c) Aggregate R in an empty backup L;
      (d) Aggregate R in a protected primary L that is not empty. For that condition and the next one below, L can be an 1:1 or 1:N primary L, but not an 1 + 1 primary L;
      (e) Aggregate R in a protected primary L that is empty;
    - if R is 1 + 1
      (a) Aggregate R in an 1 + 1 primary L that is not empty;
      (b) Aggregate R in an 1 + 1 primary L that is empty;
    - if R is 1:1
      (a) Aggregate R in an 1:1 primary L that is not empty;
      (b) Aggregate R in an 1:1 primary L that is empty;
      (c) Aggregate R in an 1:1 primary L if the removal of one or more LP LSPs of L releases enough bandwidth to install R;
    - if R is 1:N
      (a) Aggregate R in an 1:N primary L that is not empty;
      (b) Aggregate R in an 1:N primary L that is empty. For this condition the following rule needs to be accomplished: Let k be equals to the N primaries protected by the backup of L. Then the arithmetic mean of the sharing index among these k lightpaths has to be lower than the mean of any other different k lightpaths. The sharing index

of L indicates the sharing percentage of the fiber with the other (k−L) lightpaths;

(c) Aggregate R in an 1:N primary L if the removal of one or more LP LSPs of L releases enough bandwidth to install R;

- *Policy Group 3 (G3)*: Basically, this group of policies performs the same tasks as the G2. However, there are two main differences. The first one is; if the level of protection required by the request is not available, this group tries to aggregate the flow in a lightpath with a higher level of protection (if there is one available). This approach is specifically used for 1:N and, as a consequence, the 1:N request can be accommodated in an 1:1 lightpath. The second difference is that this group allows a given 1:N group to be broken in order to attend 1:1 requests. Thus, when an 1:1 request arrives and there is no such a level of protection to attend the flow, the policy breaks an 1:N group (if there is one available) so that one of the primary lightpaths of the 1:N group becomes the primary lightpath of the 1:1 level of protection. The backup lightpath of the 1:N becomes the backup of the 1:1 protection. The remaining N−1 primary lightpaths become unprotected. Note that these two differences are inversely related.

5.2.1 Results of the Second Group

To simulate the policies of the first group, it was not necessary to have a physical topology. Only the lightpaths connecting the two border nodes were enough to perform the simulations. However, the policies for failures are more complex and need a scenario with a physical topology. The physical topology is necessary to randomly simulate a failure in a fiber. By failing a fiber, we are able to determine which lightpaths are inside the fiber and then collect data related to the flows that are aggregated within each lightpath. In a previous work [4], we simulated the policies with a very simple physical topology. In this current work, the simulations are done over the NSFNet network that represents a real network topology.

The physical topology of the NSFNet network is shown in Fig. 11. The lightpaths are created from node 0 to node 12 following different routes. Each physical link has two unidirectional fibers (one for each direction) and each fiber has 10 lambdas (wavelengths) of 1 Gb/s. With that topology it was possible to create 36 (36 Gb/s) lightpaths between the ingress and egress nodes. Specifically for our simulations representing the above scenario, we created four unprotected lightpaths, six 1:N, two 1:1 and two 1 + 1. For the 1:N scheme of protection it was defined that exists one backup lightpath for three primary lightpaths (protection 1:3), resulting in six groups of 1:3 (6 * (1 + 3) = 24 lightpaths). For 1:1 and 1 + 1 schemes of protection, there exists one backup lightpath for each primary lightpath. Summarizing, 36 lightpaths were created: 24 (1:N) + 4 (1:1) + 4 (1 + 1) + 4 (unprotected). Other scenarios can be easily created by modifying the quantity of lambdas in each fiber, the bandwidth of each lambda or the type of protection.

In order to validate the policies, eight different traffic loads were injected in the network, from 80% (0.8) to 200% (2.0) of the bandwidth of the network (36 Gb/s).

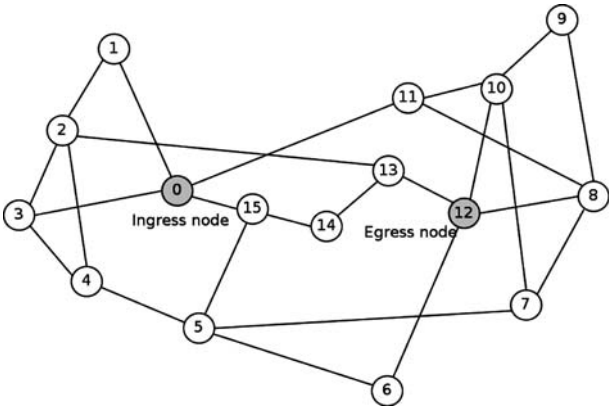

Fig. 11 NSFNet network topology

With these loads it was possible to evaluate the behavior of the policies in scenarios where the generated amount of traffic is lower than the capacity of the network and to the other extreme, we stressed the network with a high load. The percentage of generated traffic flow for each request is as follows: 35% for unprotected, 15% for 1:N, 20% for 1 + 1 and 30% for 1:1. These traffic flows were generated taking into account the percentage of the network load. For example, the amount of requests generated for the 120% load (1.2) is approximately: 36 Gb/s (capacity of the network) * 1,2 (traffic load) * 0,3 (1:1 percentage) = 13 Gb/s. The minimum bandwidth of a request is 50 Mb/s and maximum is 400 Mb/s. Statistically, the average bandwidth for each request is 225 Mb/s. Out of the total generated traffic, 50% is HP and 50% is LP. The results of the simulation were obtained through the average of 20 simulations. A single fiber failure is randomly generated for each iteration.

The results showed in this section consider the sequence of policies presented in Sect. 5.2 above, which was the sequence that responded with the best results.

Figure 12 shows the quantity of traffic that was admitted in the network (percentage related to the generated traffic). Note that G1 policies led to the least satisfactory results (actually such group of policies is the simplest one). The G3, considered the most sophisticated group, performs better when compared to the two other groups. It is important to point out that the percentage of admission depends on how the requests are aggregated within each lightpath.

Figure 13 depicts how much of the admitted traffic is HP. This is plotted for each group of policies. The G3 is the group that admitted the largest amount of HP traffic. G2 and G3 have almost the same results until the load of 1.2. However, G3 presented better numbers with higher loads. Differently, the G1 admitted around 47% with the load of 0.8, decreasing gradually reaching to 36% with load of 2.0. Observing the Fig. 13, it is possible to verify that the policies defined in G3 prioritize the HP traffic during the admission process.

Figure 14 presents the quantity of flows admitted specifically for 1:1 traffic. While G1 and G2 admitted almost the same percentage of traffic, the G3 group

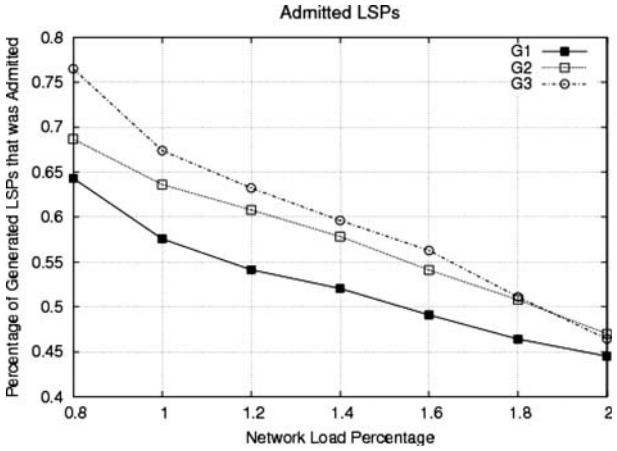

Fig. 12 Percentage of admitted traffic

admitted a higher number of flows. As an example, G1 and G2, with the load of 0.8, admitted about 14%, and G3 admitted 27%. This difference continues until the load of 2.0. Note that the good performance of G3 is because it is capable of breaking 1:N lightpaths groups in order to admit 1:1 traffic. Since more 1:1 traffic was generated in the simulations, the G3 proved to be efficient for this kind of scenario. The G3 is indicated for scenarios where the network has a considerable quantity of 1:N lightpaths and the quantity of IP/MPLS flows with protection 1:1 is higher than the network can admit with the available 1:1 lightpaths.

Figure 15 depicts the percentage of failed HP LSPs after the occurrence of a failure. In order to generate this figure, a single and randomly fiber failure was

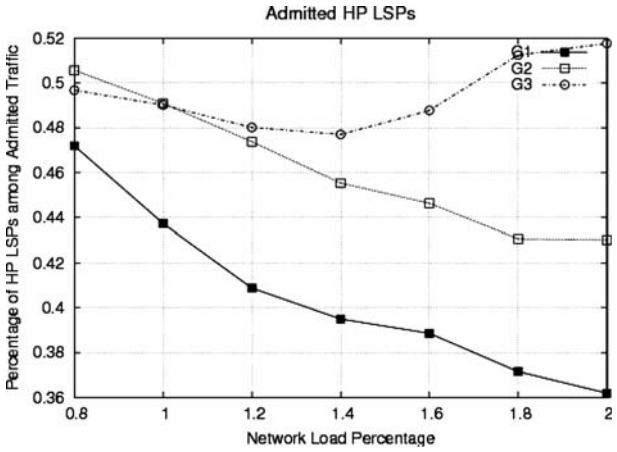

Fig. 13 Percentage of HP admitted traffic

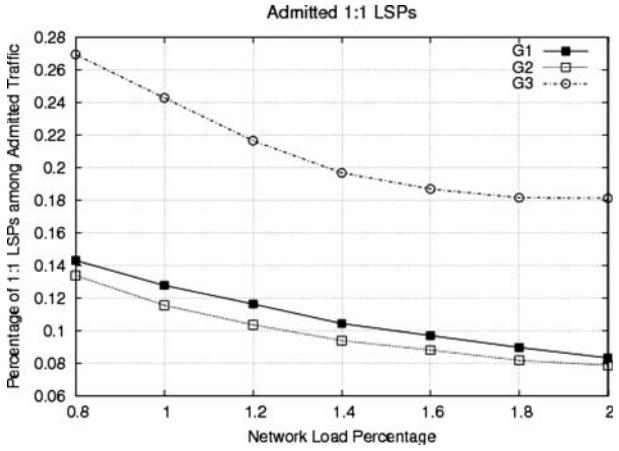

Fig. 14 Percentage of 1:1 admitted traffic

generated and the quantity of HP LSPs that were located within that fiber was counted. Figure 15 shows that G1 performs better than both G2 and G3 for all traffic loads, except for those lower than 1.0. This event occurs because the number of HP LSPs admitted using G1 is smaller than that of HP traffic admitted using G2 and G3 (see Fig. 13). Although the G3 group has admitted more flows, the good performance of such group of policies is verified with traffic loads lower than 1.0.

Figure 16 shows the percentage of LSPs that were blocked after the event of a failure. In this scenario, the system tries to readmit the failed traffic, and the LSPs that could not be readmitted are blocked. The goal here is to show that G3 is capable of readmitting more traffic than G1 and G2 for certain loads of traffic. We can see in

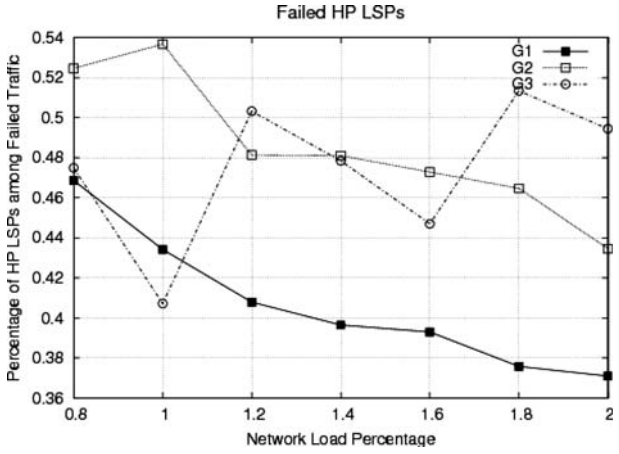

Fig. 15 Percentage of failed HP LSPs after the failure

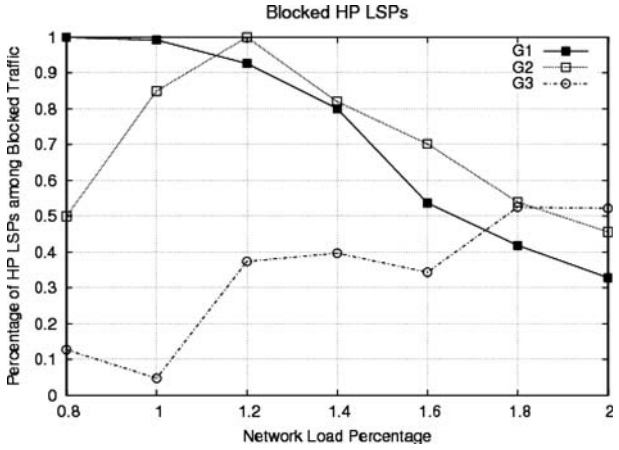

Fig. 16 Percentage of blocked HP LSPs after the failure

Fig. 16 that the number of blocked HP LSPs for G3 is lower than those for both G1 and G2 from load 0.8 to 1.6. In this interval, the G3 was able to manage and readmit HP LSPs in the network. Under higher traffic loads (more than 1.6), the G3 could not avoid the blocking due to the high number of HP LSPs in the network. However, G2 has proved to be efficient when compared to the quantity of admitted HP LSPs even in high load network scenarios as can be seen when analyzing Figs. 13 and 16.

In short, the G3 proved to be interesting in scenarios where there exists a considerable quantity of lightpaths with 1:N protection This group is indicated for a service provider that is interested in prioritizing the admission of a specific type of traffic, in this case, HP and 1:1 LSPs (break 1:N to accommodate 1:1). In addition, G3 proved to be interesting to reduce the quantity of blocked traffic after an event of failure, considering that the network load is lower than 1.8. In cases of overload (network loads above 1.8), the G2 proved to be more efficient having a lower quantity of blocked traffic.

We envisage that each provider could define meta-policies to drive the usage of the policies presented above. Depending on the traffic matrix, traffic forecasting and the network capacity, each domain could be able to apply the best set of policies by combining them to satisfy local constraints. This could be done on the fly by analyzing the current state of the network and also by having some information related to the future traffic. This discussion is left for further work.

To test the defined policies, we developed a simulator using the Java language. The policies are stored in an *ArrayList* that acts as our Policy Repository. The NSFNet network is represented as a graph in the simulator and the Dijkstra algorithm was used to calculate the paths to establish lightpaths between the two chosen nodes. The first-fit algorithm was used to select the lambda in each fiber. The first-fit chooses the first available lambda in a fiber. After the lightpaths are created, the requests are generated and sequentially sent to the AC which in turn is responsible for dispatching each request to the PM to apply the policies.

6 Conclusions

We have presented an architecture to manage the integration between the IP/MPLS layer and the optical network layer taking into account two different classes of services: HP LSPs and LP LSPs. We defined two groups of policies. The first one encompasses policies intended to reduce the number of LP LSP removals and consequently the necessity to reroute the LP LSPs in the optical layer. The second group of policies aggregates IP/MPLS flows within lightpaths in a way that when a failure takes place, the impact of such failure is reduced. Many scenarios were tested and the results were shown for both groups. For the first group, such results indicated that the policies proposed in this paper seem to be appropriated once the number of removals by applying them is significantly reduced. For the second group, the policies work with the idea that optical networks have a high amount of available bandwidth in each physical link. If such a link comes to fail, the quantity of data that will be lost is consequently very high. Solutions that are only based on schemes of protection such as 1 + 1, 1:1 and 1:N have been widely discussed. Such solutions can be improved if the type of traffic being transported within a lightpath is considered when aggregating the flows. The policies defined in this paper showed that the number of IP/MPLS flows that are blocked when applying the policies is smaller when compared to a scenario that does not use the policies.

Furthermore, the relation between the quantity of bandwidth of the network and the percentage of LP LSPs and HP LSPs used to estimate an upper bound in terms of quantity of traffic is a useful parameter to support the transport network provider on deciding when the usage of policies is recommended.

The policies defined in this work were implemented in an architecture that is composed of an AC, a PM, a FM and a RM. The efficient usage of the optical resources should be the primary goal to mitigate future investments and increase the profits in the network business. We strongly believe that the policies defined here can act as a starting point for providers to boost the usage of the optical network resources. In addition, new classes of traffic may be defined and the idea of having meta-policies could also be considered.

Acknowledgments The authors would like to thank CNPq, CAPES, FAPESP and Ericsson Brazil for their support.

References

1. Mannie, E.: Generalized Multi-Protocol Label Switching Architecture. IETF RFC 3945, October 2004 (2004)
2. Verdi, F.L., Madeira, E., Magalhães, M.: Policy-Based Admission Control in GMPLS Optical Networks. First IEEE International Conference on Broadband Networks—Broadnets'04 (formerly OptiComm), pp. 337–339. San Jose, USA (2004)
3. Verdi , F.L., Carvalho, C., Magalhães, M., Madeira, E.: Policy-Based Grooming in Optical Networks. Fourth IEEE Latin American Network Operations and Management Symposium (LANOMS 2005), pp. 125–136. (2005)

4. Carvalho, C., Verdi, F.L., Madeira, E., Magalhães, M.: Policy-based Fault Management for Integrating IP over Optical Networks. The 5th IEEE International Workshop on IP Operations & Management (IPOM'05), vol. 3751, pp. 88–97. LNCS-Springer-Verlag (2005)
5. Iovanna, P., Setembre M., Sabella R.: A traffic engineering system for multilayer networks based on the GMPLS paradigm. IEEE Netw. **17**(2), 28–37 (2003)
6. Sabella, R., Settembre, M., Oriolo, G., Razza, F., Ferlito, F., Conte, G.: A multilayer solution for path provisioning in new-generation optical/MPLS networks. IEEE J. Lightwave Technol. **21**(5), 1141–1155 (2003)
7. Dutta, R., Rouskas, N.G.: Traffic grooming in WDM networks: past and future. IEEE Netw. **16**(6), 46–56 (2002)
8. Zhu K., Zhu H., Mukherjee B.: Traffic engineering in multigranularity heterogeneous optical WDM mesh networks through dynamic traffic grooming. IEEE Netw. **17**(2), 8–15 (2003)
9. Yao, W., Ramamurthy, B.: Dynamic Traffic Grooming using Fixed-Alternate Routing in WDM Mesh Optical Networks. International Conference on Broadband Networks (Broadnets 2004), (2004)
10. Ou, C., Zhu, K., Zhang, J., Zhu, H., Mukherjee, B.: Traffic grooming for survivable WDM networks: dedicated protection. J. Opt. Netw. **3**(1), 50–74 (2004)
11. Ramamurthy, S., Sahasrabuddhe, L., Mukherjee, B.: Survivable WDM mesh networks. J. Lightwave Technol. **21**(4), 870–883 (2003)
12. Alanqar, W., Jukan A.: A review of fault management in WDM mesh networks: basic concepts and research challenges. IEEE Netw. **18**(2), 41–48 (2004)
13. Fawaz, W., Daheb, B., Audouin, O., Du-Pond, M., Pujolle, G.: Service level agreement and provisioning in optical networks. IEEE Commun. Mag. **42**(1), 36–43 (2004)
14. OIF User Network Interface (UNI) 1.0 Signaling Specification
15. Berger, L.: Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions. RFC 3473, January 2003 (2003)
16. Vasseur, J.-F., Ayyangar, A.: Inter domain GMPLS Traffic Engineering—RSVP-TE extensions. draft-ayyangar-ccamp-inter-domain-rsvp-te-02.txt, January 2005 (2005)
17. Mannie, E., Papadimitriou, D.: Recovery (Protection and Restoration) Terminology for Generalized Multi-Protocol Label Switching (GMPLS). draft-ietf-ccamp-gmpls-recovery-terminology-05.txt, October 2004 (2004)
18. Verdi, F.L., Duarte, R., de Lacerda, F.C., Madeira, E., Cardoso, E., Magalhães, M.: Web Services-based Provisioning of Connections in GMPLS Optical Networks. The Brazilian Symposium on Computer Networks (SBRC 2005). Fortaleza, Brazil (2005)
19. ITU-T: Architecture for the Automatically Switched Optical Network (ASON), G.8080/Y.1304
20. Common Information Model, http://www.dmtf.org/standards/cim

Author Biographies

Fábio Luciano Verdi is currently a post-doc student at the Faculty of Electrical and Computer Engineering (FEEC), State University of Campinas (UNICAMP), Brazil. He received his Master degree in Computer Science and Ph.D. degree in Electrical Engineering both from UNICAMP. His main interests include computer networks, mobility, routing, service oriented architectures, inter-domain services and next generation Internet Architectures.

Cláudio Carvalho received the Bachelor degree in Computer Engineering from the Catholic University of Goiás (UCG) in 2003 and the Master degree in Computer Science from the University of Campinas (UNICAMP) in 2006. Currently, he is a Ph.D. student at UNICAMP. His research interests include policy-based management, traffic grooming, optical network management, control planes and autonomic networks.

Maurício F. Magalhães received the B.S. in Electrical Engineering from University of Brasília (UnB), Brasília, Brazil, M.S. in Automation from School of Electrical Engineering, State University of Campinas (UNICAMP), Campinas, Brazil and Dr. Engineer from *Laboratoire d'Automatique (LAG/CNRS)* and *Institut National Polytechnique de Grenoble (INPG)*, Grenoble, France. Currently he works as a Titular Professor at the School of Electrical and Computer Engineering, UNICAMP).

Edmundo R. M. Madeira is an Associate Professor at the Institute of Computing of State University of Campinas (UNICAMP), Brazil. He received both his Ph.D. in Electrical Engineering and M.Sc. in Computer Science from UNICAMP. His research interests include network management, optical networks and distributed systems. He has published over 100 papers in national and international conferences and journals. He has also supervised more than 30 master and Ph.D. students.