RESILIENT MULTILAYER TRAFFIC ENGINEERING THROUGH DYNAMIC PATH RESTORATION

Survivable data-centric automatic switched optical networks

Bart Puype, Qiang Yan, Sophie De Maesschalck, Didier Colle, Mario Pickavet, Piet Demeester
Ghent University – IMEC, Department of Information Technology

Abstract: In an IP-over-ASON scenario, the logical IP topology can be provisioned on demand by taking advantage of the flexible optical switching capabilities. Multi-layer Traffic Engineering (MTE) strategies allow to guarantee QoS and solve problems such as congestion dynamically through rerouting as well as reconfiguration of this logical topology. Network failures however can be handled as well, by letting the MTE strategy perform network recovery through dynamic path restoration. The paper discusses how a distributed MTE algorithm can be extended towards resilience, the types of traffic that need be recovered, and how it then copes with optical link failures and node failures on the optical and IP layer.

Key words: Multi-layer TE, resilience, path restoration, ASON

1. INTRODUCTION

Automatic Switched Optical Networks (ASON [1]) characteristics add a whole new layer of switching flexibility to next-generation optical networks. Implementing the User Network Interface by the OIF [2] and Generalized Multiprotocol Lambda Switching [3], these networks enable dynamic set up and tear down of end to end lightpaths through the use of optical switching equipment such as Optical Cross Connects (OXCs) or Optical Add-Drop Multiplexers (OADMs). Full optical DWDM switching becomes possible with emerging technologies such as MEMS. Traditional communications services such as voice and video are being replaced by packet based alternatives. These new multi-media applications not only require large amounts of bandwidth, but also expect a predefined Quality of Service (QoS). While traditional circuit based implementations (such as the POTS) guarantee this QoS by design, this is not the case in packet based data networks, which suffer from packet loss, delay and jitter, generally caused by congestion of processing limitations in routers and bandwidth shortage in the network topology. QoS of packet streams (such as those in IP networks), is usually handled through judicious use of Traffic Engineering (TE), for example using MPLS label switched paths.

The previously mentioned ASON however enables one to dynamically alter the logical topology that is seen by the IP (packet) layer, this by offering lightpaths over the circuit-switched physical optical topology. The process of Multilayer Traffic Engineering (MTE) combines both the classic concept of rerouting TE and this additional optical flexibility through which the logical topology can be reconfigured on-the-fly.

Regardless of whether a circuit or packet switched network is used, there will always be the issue of network faults. Unless the network resilience aspect is taken into account, occurrences such as fiber cuts or failing switches and routers will cause service interruptions or even outages. For the traditional circuit switched case, a number of survivability mechanisms exists in the form of network protection and static path restoration schemes. Both types necessitate the advance reservation of bandwidth in the form of protection or spare capacity. As discussed in [4] however, a multi-layer scenario such as IP-over-ASON will not suffice with merely network protection. Failing IP routers and optical switches may carry traffic that is only transiting these nodes and as such one can expect it to be recovered. This will however require rerouting in the IP layer, as well as setting up recovery lightpaths. This paper will explain how a MTE strategy, which under fault-free operation ensures correct behavior of the network and guarantees the QoS requirements, can be extended to cope with several fault scenarios. In these cases the MTE will act as a dynamic path restoration process. Once optical replacement bandwidth has been established, it will then proceed with rerouting traffic and perhaps further optimization of the logical topology.

In section 2, we will show how to extend an existing MTE strategy to allow it to cope with network failures. 2.1 discusses how these failures can act as triggers for the traffic engineering process, while 2.2 examines how the type of affected traffic and failure influences the possible actions that can be taken.

We will further illustrate these concepts with a case study performed on a resilient MTE strategy in section 3, based on a proactive MTE design previously discussed in [5].

2. EXTENDING MULTILAYER TRAFFIC ENGINEERING TOWARDS SURVIVABILITY

A typical MTE strategy consists of a number of recurring aspects. It needs a routing algorithm which directs traffic flows over the logical topology, using for example electrical LSPs. It also has an algorithm that configures the logical topology. Both these processes will interact with each other. Reconfiguring the logical topology will likely change routing costs, and the way traffic is routed has a direct influence on parameters such bandwidth usage, perceived QoS, congestion, which in turn act as input towards the MTE strategy. Exactly how an MTE action is triggered remains an open issue here (see [5],[6] for more information), but surely the main driving force behind MTE actions under normal operation is the continuously evolving traffic demand that is applied to the logical topology.

2.1 Network failures as MTE triggers

As previously mentioned, optical networks usually incorporate network protection (or restoration) schemes, usually reserving some kind of backup capacity such as the protection path A–B–C in *Figure 1* (a).

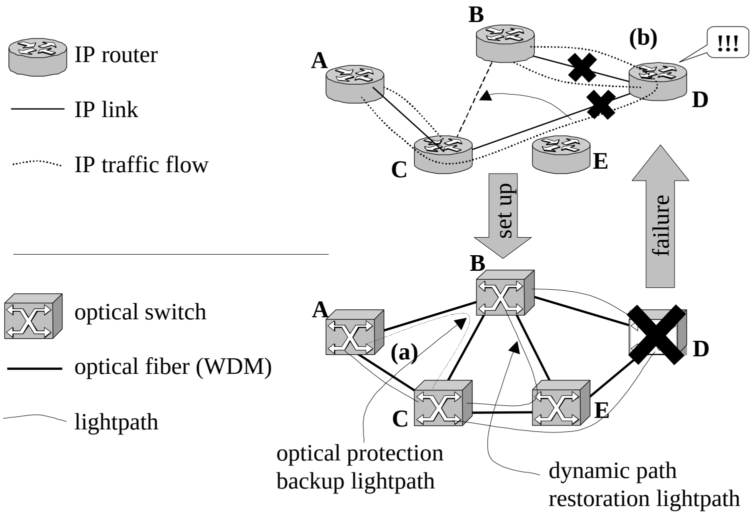


Figure 1. Protection and dynamic path restoration mechanisms

While it is certainly possible to bypass these and provide full survivability through client-layer resilience, one is more likely to prefer recovery mechanism in the optical layer, since these can act faster and also more efficient because of larger granularity – and thus impact QoS less.

However, not all network failures can be recovered in the underlying optical network. Therefore survivability mechanisms in both client (IP) and server (optical) layer must be coordinated, such that IP recovery mechanisms do not interfere when optical protection or restoration is possible. One way to establish this is to introduce a hold-off timer [7], which delays client-layer mechanisms for a given amount of time, anticipating optical recovery.

Survivability through MTE is equivalent with client-layer mechanisms in this discussion. The routing part of the MTE process acts from the client-layer, as does the reconfiguration part, since it needs to integrate with the routing, and typically signals lightpath set up or tear down through the UNI.

Once it is well-defined when to trigger client-layer recovery, the network failure can be passed on to the MTE strategy, as shown in *Figure 1* (b), for an optical node failure at D. This will remove IP links B–D and C–D from the logical topology, causing the traffic following A–C–D–B to need recovery. This is done by setting up new bandwidth between B–C, which needs to be signaled to the optical layer.

Although it will usually deal with changing traffic patterns, recovering from a fault scenario can be handled exactly the same way, the only difference being that besides a possible change in traffic, a failure also entails a change in the logical topology. This is not a problem for the MTE strategy: a network failure simply leads to an discrepancy between traffic demand and logical topology, just as if one would apply a change in traffic to an existing topology.

Since an optical network failure affects many lightpaths (i.e. a large part of the logical topology), an important condition is that the MTE strategy can cope with large changes in demand (or topology), more specifically, that it remains stable and does not suffer from slow convergence in these cases. A related issue is the fact that a failure affects many lightpaths at once, meaning one will see a burst in signaling traffic, which could adversely affect the MTE strategy, since it too relies on signaling to perform its actions.

2.2 Types of network failures and affected traffic

The impact a network failure has on the network, and the extent of the MTE actions required to take, depend on the exact type of the failure. Mostly because different failures will affect different types of traffic, since the IP layer traffic (and whether it can be routed over the logical topology) is the main concern of the MTE algorithm.

2.2.1 Types of affected traffic transiting nodes

When looking at a single node, it’s possible to separate three different classes of traffic in a multi-layer IP-over-ASON scenario. This node combines both an optical switch and an IP router (*Figure 2*).

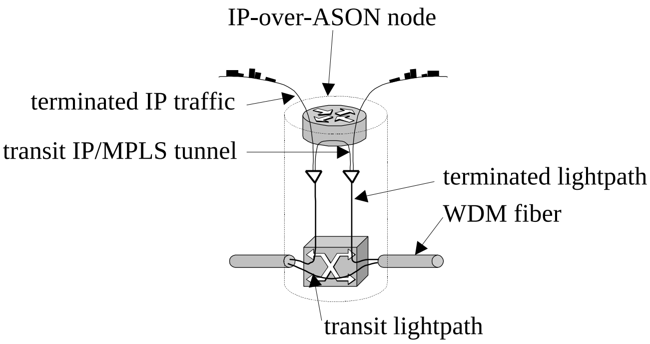


Figure 2. IP-over-ASON node and types of traffic

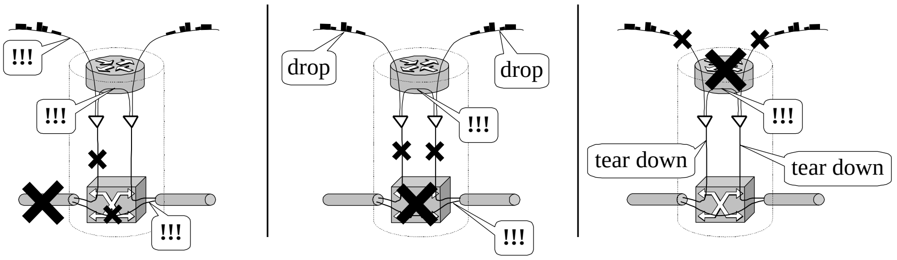
Firstly, there is traffic that does not leave the optical switching matrix. This consists of any traffic that is carried by lightpaths that are switched – but not terminated – at the optical node. This type can be fully recovered by optical protection or restoration using pre-allocated backup paths.

However, some lightpaths will be terminated at the node. Some of the IP traffic carried by these paths will be destined for the attached IP router. Optical recovery in case one of these lightpaths is affected, is still possible, and will be performed by switching to an optical backup path at this node.

To optimize optical bandwidth usage, the MTE strategy will still allow multi-hop IP traffic over the logical topology, so consequently, a third type of traffic, which is also optically terminated, is merely transiting the IP router. This kind of transit traffic will be forwarded using MPLS tunnels, specified by the (IP part of the) MTE algorithm. For this traffic, whether or not one needs to take MTE actions depends on the exact type of network failure.

2.2.2 Types of network failures

Fiber cuts are the most frequent optical network faults (*Figure 3*, a). Optical network protection or possibly restoration is the preferential mechanisms to cope with these, especially in DWDM networks where a large number of lightpaths are multiplexed into a single fiber. Optical survivability schemes allow to recover these with minimal signaling overhead and downtime. In an ASON network, necessary additional spare capacity will be allocated at the time a lightpath is requested, to provide for these resilience mechanisms. As far as the MTE algorithm is concerned, no failure has happened. Note though that in the case of optical restoration, recovery may fail due to blocking of recovery lightpaths (because of limited optical capacity), in this case the link failure may be escalated to the client layer (i.e. to the MTE strategy), to see if rerouting or reconfiguration can help instead.



(a) single fiber failure (b) optical node failure (c) IP router failure

Figure 3. Types of network failures

The other, and in this discussion more interesting network failures concern complete node failures on either the IP or optical layer. In the case of an optical node failure (*Figure 3*, b), the attached IP router gets isolated from the rest of the network. All lightpaths that both transit and terminate at the node, go down. Traffic terminated at the IP layer can not be recovered – assuming the IP router does not have connectivity to a backup optical switch available. However, transit traffic in both the optical layer (switched lightpaths) and IP layer (forwarded MPLS traffic flows) can and needs to be recovered. To do this, the affected traffic needs to be rerouted, and the logical topology will likely need to be repaired, although the isolation of the failing node will remove some traffic, thereby unloading somewhat the IP layer. On the other hand, the removal of an optical node removes a lot of bandwidth from the optical layer

In the other case of a router failure in the IP layer (*Figure 3*, c), no lightpaths are affected. The only traffic that needs to be recovered, will be the transit traffic in the IP layer, again rerouting will be necessary here. Of course, since the optical switch is still functional, recovery lightpaths can still be set up over the node. Also, since lightpaths terminated at the node have now become useless, it may be beneficial during prolonged IP router failures to tear these down to free optical capacity.

3. CASE STUDY

In order to illustrate the previous discussion, we will now present a case study showing some results from simulations performed on a resilient multi-layer traffic engineering strategy. The MTE algorithm is based on a proactive strategy described in more detail in [5].

3.1 MTE strategy

The resilient MTE strategy used in the case study is depicted in *Figure 4*. Since it is a proactive algorithm, it continuously recalculates routing and configuration of the logical topology. Whenever a significant optimization is found, a MTE actions is performed, meaning the new routing or configuration is signaled to the network.

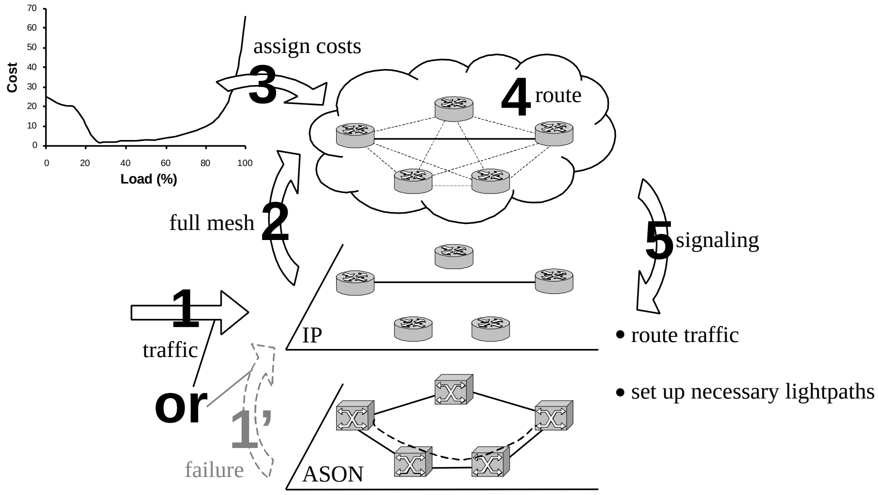


Figure 4. Resilient MTE process

The algorithm simulates routing in a virtual full mesh, using a well designed cost function (also shown in the figure). IP layer traffic is assumed to be tunneled in MPLS Label Switched Paths (LSPs), so that IP layer TE is possible, and routing considers discrete traffic flows. It will make sure to match the actual IP layer routing with the routing in the virtual full mesh, however only setting up IP links (lightpaths) when really necessary (i.e. when occupied by at least one traffic flow). The amount of used lightpaths is furthermore minimized through the IP routing cost function, which has a higher cost for lightly loaded IP links (thus avoiding the usage of a large number of such lightly loaded links). We assume there is enough capacity available in the optical layer to set up these lightpaths.

Under a constant traffic pattern, the algorithm will converge towards an optimal logical topology. The normal MTE process however entails measuring traffic demand at regular intervals, and applying this as input (1) to the strategy, retriggering the optimization process (see *Figure 4*). Next step would be to simulate a full mesh (2), and assign IP link costs in this full mesh (3). Note that the cost depends on the load of the IP link, so this is where the traffic demand input comes into play. Next, the traffic flows are routed (4), and after that necessary signaling is performed (5), in which routing tables and MLSP forwarding tables are adjusted and lightpaths are set up or torn down if needed.

Since the decision on what MTE actions to take is made mainly inside the full mesh simulation, it is possible to extend this MTE strategy towards resilience (it assumes full logical topology flexibility). Instead of applying new traffic measurements (1), we now signal the MTE process with an optical or possibly an IP-layer failure (1’). Note that in a proactive process, where traffic flows are routed only periodically, one will prefer to give priority to a failure, meaning the traffic flows affected by the failure will be rerouted immediately; thus making sure service interruptions remain minimal.

Furthermore, some failures will modify the traffic demand: when a node goes down, all traffic terminated there will be removed from the network, causing the removal of the related traffic flows. Again the MTE algorithm will not wait for the next traffic measurement to take this knowledge into account while rerouting the remaining affected flows that do need to be recovered (i.e. transit flows, in either the optical or IP layer).

3.2 Used topology

For the network topology and the distribution of optical switching and IP routing equipment, we assume each IP node has an optical OXC associated with it. The physical optical topology itself is presented in *Figure 5*.

It represents an Italian backbone network. The topology is such that the node in Rome acts as a central point for many lightpaths, in order to see the large impact a node failure there would have. We have chosen the optical capacity so that there is ample optical bandwidth available for this case study – the total number of lightpaths over the optical topology is typically 50.

The optical routing and wavelength assignment of lightpaths in this case study, is not handled by the MTE strategy itself, but directly by the optical layer. This means the optical layer is seen by the MTE algorithm as a blackbox, that delivers lightpaths between any two IP routers. We have utilized simple shortest path routing and assumed non wavelength translating nodes for this optical layer.

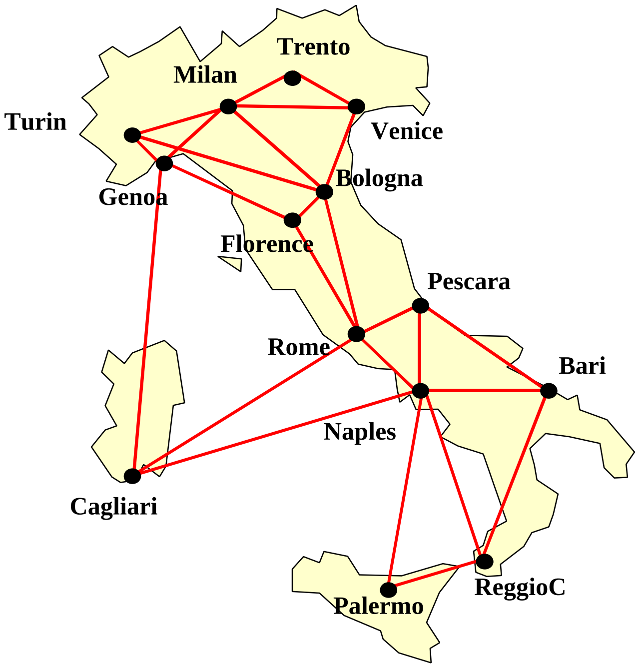


Figure 5. Physical optical topology

3.3 Results

The results listed below concern several network failures under a typical traffic pattern, which is depicted in *Figure 6* – we show IP load for each optical link, in percentage of the bandwidth of a lightpath. Note that a single lightpath circuit is not fully utilized by packet traffic, in order to avoid packet loss due to IP routing buffer overload. Typically, 80% utilization of a single circuit (lightpath) will be the maximum allowed, while our usage of a MTE strategy may lower this even more in order to guarantee a certain QoS. This means a total IP load of for example 300% over an optical link will require more than three lightpaths. The traffic pattern here requires a total of 50 lightpaths to be set up over the optical topology.

The three types of failures mentioned before will be discussed separately. For each type, all possible single failure are examined (e.g. for link failures, all single link failures are simulated).

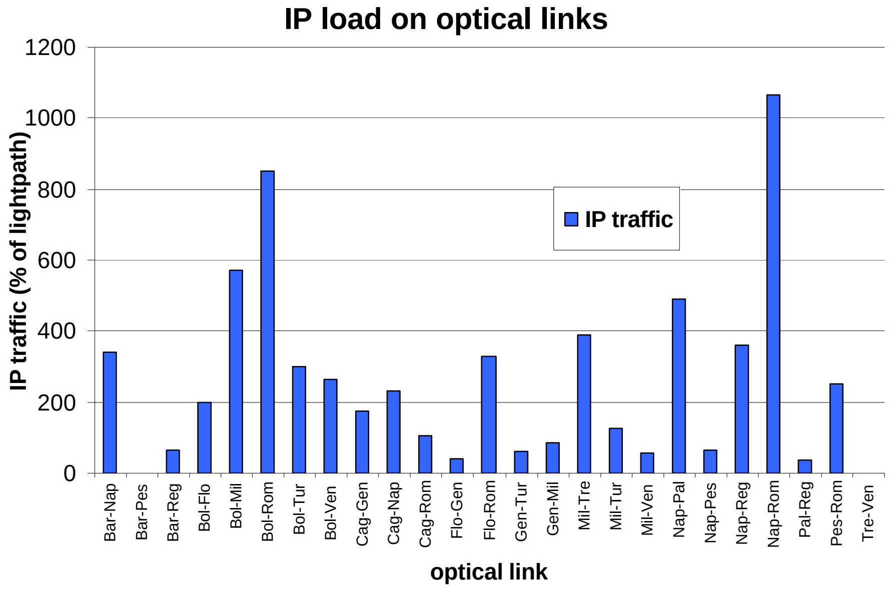


Figure 6. Used traffic pattern (IP load)

3.3.1 Optical link failures

First, we will introduce optical link failures into the network. Although fiber cuts are usually handled by the optical layer through network protection mechanisms, these results are nevertheless interesting. Since a fiber cut does not change the traffic demand (no IP routers becomes isolated), we can see how the MTE strategy will behave for a purely survivability related situation.

Figure 7 shows the effect of a single fiber cut on the optical layer. For each fiber cut, but for the same traffic pattern, we show both unaffected lightpaths (i.e. not routed over the failing link) and recovery lightpaths. As can be expected, the amount of replacement lightpaths scales with the IP traffic over the failing optical link. Compare the results for links Bol-Rom and Nap-Rom with the traffic pattern in *Figure 6*, for example.

The initial amount of lightpaths (50) is compared with the number of set up lightpaths after the recovery converges (*Figure 8*): the load on the optical topology is lower afterwards, in most cases. This is in part because a certain amount of inertia is built into the MTE strategy [5], in order to prevent long convergence times. Small changes – such as the link failures here – will retrigger the MTE strategy, usually improving the logical topology.

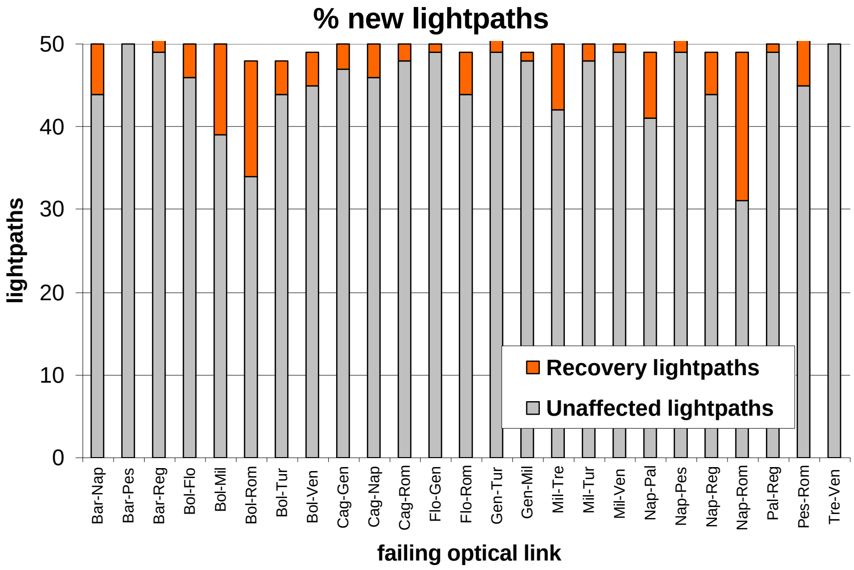


Figure 7. Affected lightpaths for each failing optical link

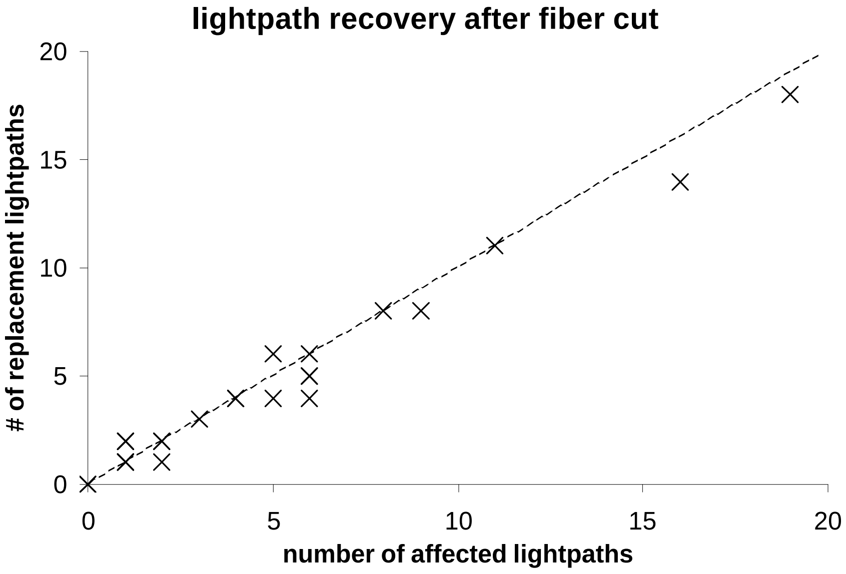


Figure 8. MTE link failure recovery performance

3.3.2 Optical node failures

Node failures have a larger impact on the network. Also, since some traffic is dropped from the demand, we see both recovery and traffic adaptation aspects in the MTE strategy. On *Figure 9*, we have the same traffic pattern as in *Figure 6*, but now the amount and type of transit traffic for each node is shown (also the amount of terminated traffic, which is fairly equal for all nodes). Similarly to the larger bandwidth usage on links Bol-Rom and Nap-Rom, we can see high amounts of transit traffic for the four nodes with a high degree, which make up the central axis of the physical topology (Milan, Rome, Bologna and Naples).

Also note that in most cases, the amount of IP layer transit traffic remains below 100% (of the bandwidth of a lightpath). The remaining amount of transit traffic is shifted to the optical layer in switched lightpaths. This is of course the exact goal of the MTE strategy: lowering the load in the IP layer by taking advantage of the higher granularity switching capabilities of the optical layer.

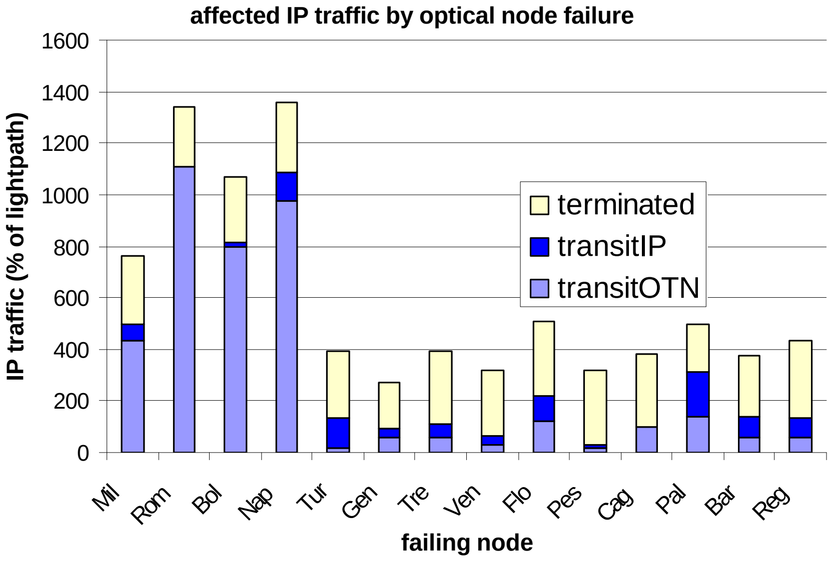


Figure 9. Type and amount of transit traffic in each node

Now, a failing optical node will take out all optical links terminated on it, and isolate the IP-router. This means all lightpaths either terminated on or switched by the node will be removed from the network, as the MTE strategy tears these down to free optical bandwidth in the remaining physical topology.

Again the number of unaffected and recovery lightpaths is shown in *Figure 10*, this time for each (single) optical node failure. As expected, a failure in one of the four important nodes mentioned before, causes a lot of optical reconfiguration. In the other cases, mostly IP transit traffic needs to be rerouted, which may or may not require replacement lightpaths, since the MTE strategy can redirect affected traffic entirely within the IP layer (some IP layer bandwidth may have been freed as well).

The final amount of required optical bandwidth after convergence depends only slightly on the exact optical node that failed (required lightpaths range from 40 to 44), this is again because the total dropped (IP) traffic (i.e. terminated at the failing node) is fairly constant over all nodes.

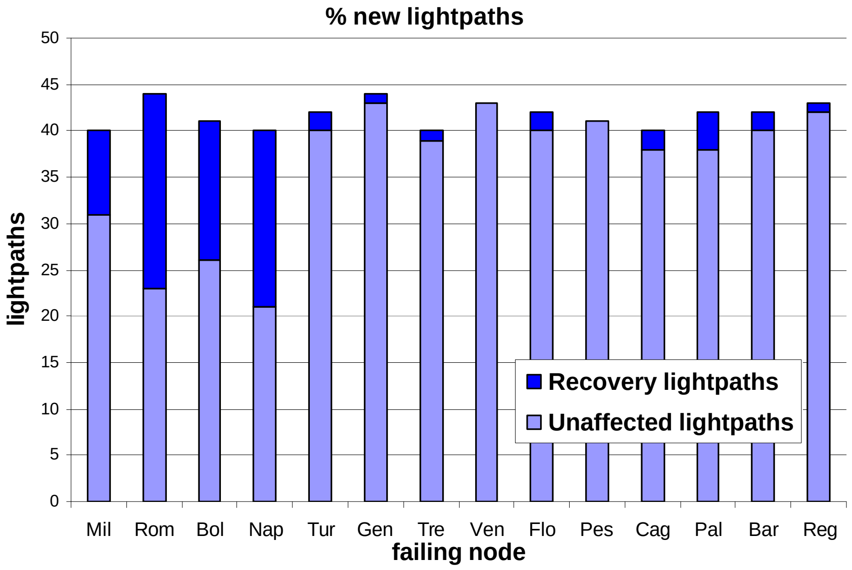


Figure 10. Affected lightpaths for optical node failures

Figure 11 shows a plot of the number of required new lightpaths against the amount of total affected IP transit traffic (sum of IP and optical layer transit traffic). A ratio affected IP bandwidth / required lightpath can be derived from all simulations on this scenario, and we see that IP links (lightpaths) are, on average, loaded at about 55 % of their capacity. This is to optimize QoS parameters. This average load can be adjusted by modifying the MTE cost function (see *Figure 4*), more specifically by shifting the cost minimum towards higher IP link loads [5]. However, this shows that the performance of the MTE strategy for the recovered traffic is not affected by the fact that routing and logical topology configuration were triggered by a network failure, but is rather comparable to the normal operation triggered by changing traffic patterns.

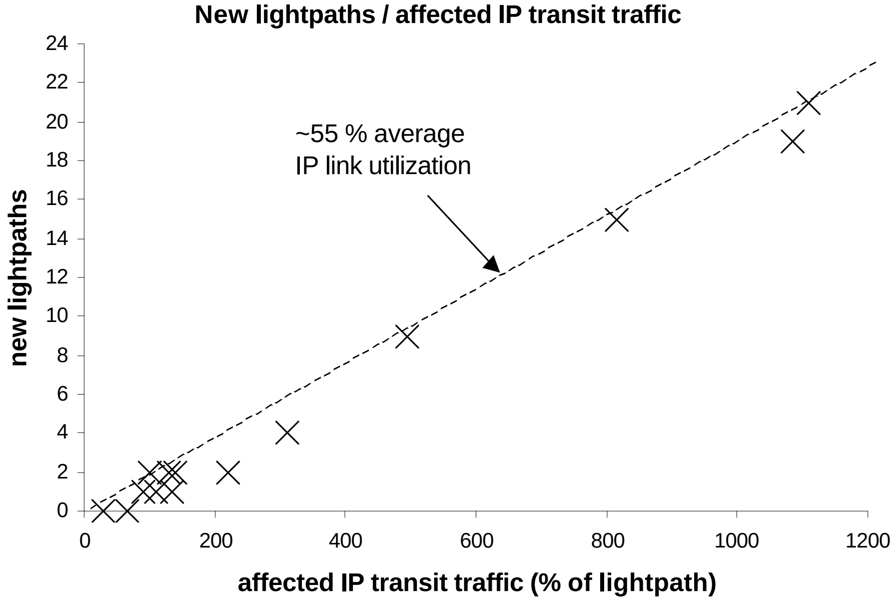


Figure 11. Amount of new lightpaths for affected transit traffic

3.3.3 IP router failures

In the case of an IP router failure, no optical layer lightpaths are affected, although any lightpaths terminated at the optical switch associated with the failing IP router, now become useless. The major task of the MTE strategy here will then be freeing the optical capacity taken by these unused lightpaths.

The same simulations as for the optical node failure case are now shown for IP router failures in *Figure 12*. Since only the IP layer transit traffic has to be recovered, one usually suffices with IP layer rerouting only. In the few cases where actual new lightpaths are set up, this is caused by either the logical topology not having enough free IP bandwidth, or the IP router failure triggering a new reconfiguration sequence, leading to some possible optimizations when replacing a few working lightpaths.

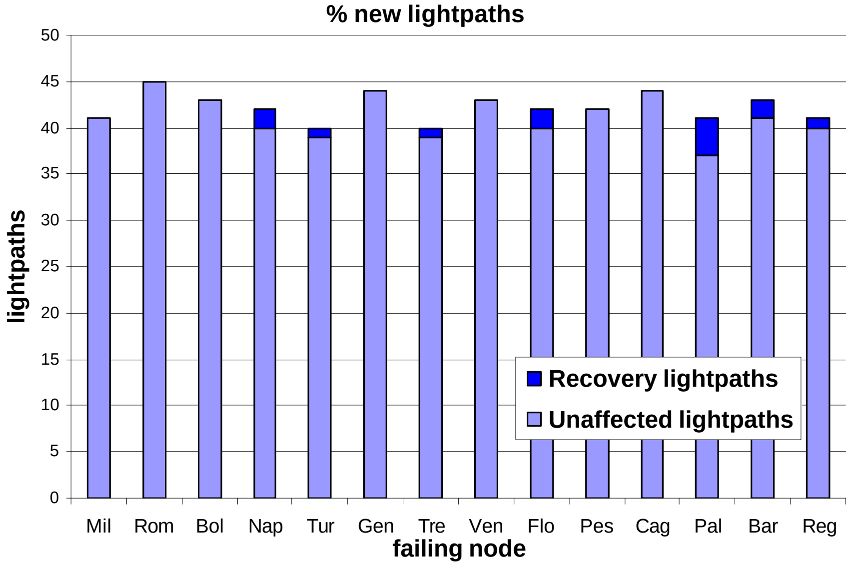


Figure 12. Affected lightpaths for IP router failure

4. CONCLUSIONS

We have discussed how an existing MTE strategy can be extended with survivability aspects in an IP-over ASON scenario. This way it can cope with network failures that have been escalated from the optical layer to the IP layer, as some failures cannot suffice with optical protection (or restoration) alone.

The type of network failure, and more importantly the type of traffic (terminated or transit at the failure) that is affected by it, dictate for a large part which MTE actions are required.

We have presented results for optical link failures, but more importantly also for optical node as well as IP router failures, as these cause large damage to the logical topology and have an influence on the traffic demand at the same time. The MTE strategy has shown to be sufficiently robust to reroute affected traffic if needed, setting up new bandwidth over the working part of the optical topology, and possibly tearing down lightpaths that traverse the failure, thereby freeing additional optical capacity to aid in the recovery. Additionally, the MTE performance has proven not to be affected when the MTE process is triggered by network failures, compared to the normal operation adapting to changing traffic patterns.

ACKNOWLEDGEMENTS

Part of this work has been supported by the IWT and ITEA through the Transparent Backbone Optical NEtwork Simulator (TBONES) project; and by the IWT through the GBOU project Optical Networking and Node Architectures, and also through the post-doctoral scholarship granted to Didier Colle.

REFERENCES

1. ITU-T Rec. G.8080, “Architecture for the automatically switched optical network (ASON)”, November 2001
2. OIF2000.125.5 , “User Network Interface (UNI) 1.0 Signaling Specification”, June 2001
3. E. Mannie et al., “Generalized Multi-protocol Label Switching (GMPLS) Architecture”, Internet Draft, Network Working Group, 2002
4. S. De Maesschalck et al., “Intelligent Optical Networking for Multilayer Survivability”, IEEE Communications Magazine, feature topic on “Resilience in Communication Networks”, Vol. 40, No. 1, January 2002, pp. 42-49
5. B. Puype et al., “Multi-layer Traffic Engineering in Data-centric Optical Networks, Illustration of Concepts and Benefits”, COST266/IST Optimist workshop – ONDM 2003, Budapest (2003), pp. 211-226
6. Qiang Yan et al., “Influence of the Observation Window Size on the Performance of Multi-layer Traffic Engineering”, ITCOM, Florida (2003), pp. 203-214
7. D. Colle et al., “Cost-efficient Deployment of Survivable Next-generation IP-over-Optical Networks”, ITCOM, Florida (2003), pp. 215-229