Optical cost metrics in Multi-layer Traffic Engineering for IP-over-Optical networks

Bart Puype[*], **Qiang Yan**[*], **Sophie De Maesschalck**[*], **Didier Colle**[*], **Kris Steenhaut**[**],
Mario Pickavet[*], **Ann Nowé**[**], **Piet Demeester**[*]
[*]*INTEC, Gent University – IMEC, Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium*
[**]*Vrije Universiteit Brussel, Pleinlaan 2, B-1000 Brussel, Belgium*
[*]*Tel: +32-9/264.99.86, Fax: +32-9/264.99.60, e-mail: bart.puype@intec.UGent.be*

ABSTRACT

Automatic Switched Optical Networks (ASONs) or Intelligent Optical Networks (IONs) facilitate setting up or tearing down end-to-end lightpaths dynamically. For an IP-over-Optical network scenario, Multi-layer Traffic Engineering (MTE) allows to use this optical flexibility to reconfigure the logical IP topology in order to cope with changing traffic demands. Utilising an overlay model limits the information exchange between both layers. However, we show that by incorporating even simple optical cost metrics in the MTE process, we can attain significant cost savings in the optical layer, compared to the case where no information (beyond reachability) is communicated - and this with minimal impact on IP layer performance.

Keywords: Multi-layer Traffic Engineering, Automatic Switched Optical Network (ASON), cost metrics, routing

1. INTRODUCTION

The continually increasing bandwidth volume over data – and in particular – IP-based networks, has traditionally been transported over point-to-point optical fibre links. Using DWDM technology, optical technology offers Tbit/s bandwidths, at 2.5 or 10 Gbit/s per optical wavelength channel. The advent of new switching technologies as well as protocols such as MEMS [1] and GMPLS [2] has given rise to new flexibility in the optical layer, leading to Automatic Switched Optical Networks [3]. ASONs offer an optical layer which can dynamically set up end-to-end optical connections (lightpaths) for a variety of client-layer services. In the case of an IP-over-Optical scenario, these lightpaths become the IP links in the IP logical topology. In GMPLS terminology, these lightpaths become optical LSPs, analogous to traditional electrical LSPs (which allow traffic engineering strategies to decouple routing and forwarding), albeit with a much larger bandwidth granularity.

Multi-layer Traffic Engineering (MTE) extends TE with cross-layer techniques using this flexibility in the underlying optical layer. MTE encompasses Single-layer TE (i.e. routing) in both the IP and optical layer, as well as the interworking between the two, namely the dynamic configuration of the IP logical topology by setting up and tearing down optical bandwidth. Due to the large granularity of lightpaths, MTE has to solve a dynamic grooming problem in the IP layer.

An integrated model allows access to complete knowledge of both IP and optical layers, in terms of IP router capacities, WDM topology, optical switch (OXC) characteristics, etc. This enables one to perform intelligent routing and multi-layer grooming of incoming traffic (e.g. [4]). An overlay model on the opposite, separates IP and optical layer so that information exchange between the two is limited. In this paper, we show that by incorporating even simple optical cost metrics in an IP-situated MTE process, it is possible to attain significant optical cost savings, compared to the case where no information (beyond reachability) is communicated.

In section 2, we will first outline the Multi-layer Traffic Engineering strategy studied in this paper. It is presented as a MTE process which runs in the IP layer (thus managing the logical IP topology and routing, but not RWA in the optical layer), while section 3 will introduce additional optical metrics. Section 4 details a simulation study based on the concepts from this paper, illustrating the optical metrics' impact on optical resource usage and IP layer performance.

2. MULTI-LAYER TRAFFIC ENGINEERING STRATEGY

A proactive MTE strategy tries to optimise continuously both the routing in the logical IP network, as well as the logical IP topology itself. This is in contrast with a pure reactive mechanism which only responds to certain problems (e.g. a local QoS degradation). A proactive strategy will pre-emptively reroute or even reconfigure the logical topology, anticipating potential problems. Note that this entails more than intelligent routing of incoming connection requests (flows, LSPs to be set up, etc.); as some of these connections may be rerouted as a consequence of a changing logical topology. In this manner, MTE copes with changing traffic demands offered to the IP layer.

Part of this work has been supported by IST through the Integrated Project NOBEL [6], the IWT through the GBOU project Optical Networking and Node Architectures [7] and the FWO-project G.0315.04, the scholarship granted to Bart Puype and also through the post-doctoral scholarship granted to Didier Colle.

The concept of the proactive MTE strategy presented in this section is based on a cost function which is used to route traffic over the IP logical topology [5]. Additionally, the cost function (through the IP routing) determines which logical IP links (lightpaths) are to be set up.

This is done by routing at all times over a fictive full mesh IP layer. The assumption of an IP full mesh is rationalised by the fact that the underlying optical network allows us to set up any logical IP link desirable (through the use of a lightpath). The cost to use a link (in the fictive full mesh) depends on the utilisation or load of that link, or rather of the corresponding link in the actual IP logical topology – this promotes balancing the IP load over the logical topology. However, since an actual full mesh is undesirable, the routing cost will have to be such that only a number of those fictive full mesh links are used for offered traffic (the cost will also promote traffic grooming). Only this limited set of IP links will then be requested over the optical network. At any instant, there will be some IP links set up, while the remaining portion of the fictive full mesh is not available in reality.

On Fig. 1 (left part), we see the cost function used by the MTE strategy, with the load (of an IP link) on the X-axis. Also on the chart is a bar load histogram showing the number of IP links with a load in a certain load interval (for a typical logical IP routing using this cost function). The cost function is characterised by three parameters. Firstly, there is a High Load Threshold (**HLT** on the chart) – IP links with a load above HLT receive a exponentially rising cost. However, also lightly loaded links (with a load below Low Load Threshold, **LLT**) are penalised with a higher cost, which is defined by the Low/Moderate Ratio (**LMR**), indicating the ratio between cost for low loads (LC) and cost for moderate loads (MC); LMR = LC/MC. This cost penalty avoids the establishment of many and thus inefficiently used links. In other words, it is our intention to attain a network with moderately loaded links *only*, promoting grooming. The shape of the empirically designed cost function was reached using the following equation:

$$\begin{aligned} C(L) = {} & 2.\mathrm{LMR}\left[\exp\left(\frac{a(L-\mathrm{LLT})}{b}\right)+1\right]^{-1} \\ & +\mathrm{LMR}\left[\exp\left(\frac{-L}{a.b}\right)+\exp\left(\frac{L-\mathrm{HLT}}{a.b}\right)\right] \quad + \quad 1 \qquad (1) \\ & +2.\mathrm{LMR}.\exp\left(\frac{L-\mathrm{HLT}-0.15}{b}\right) \end{aligned}$$

where $C(L)$ is the cost depending on the load $L \in [0, 1]$. There are two shaping parameters $(a, b) = (4, 0.05)$, which control the curvature of the function. The first term generates the high costs for low loads below LLT = 0.20, while the fourth term is responsible for the exponential cost for overloaded links, it dominates the cost function for loads past HLT = 0.80. The second term causes the cost function to be slightly convex in the moderate load range (between LLT and HLT). LMR = 8 is the ratio between the first term for $L = 0$ and the constant third term which is always present and therefore the minimum cost in the moderate load range; in this case, 1. The highest cost is always found for overloaded IP links, because avoiding overload is more important than forcing grooming.

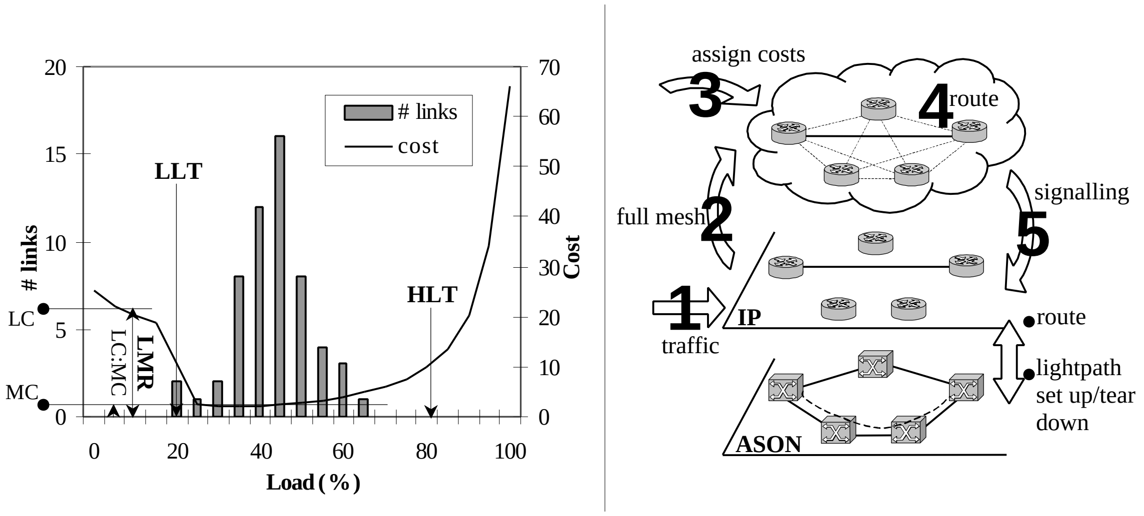


Figure 1. MTE cost function and overview of MTE process

Fig. 1 (right part) outlines the MTE strategy that is based on this cost function. We see the actual network (IP and ASON layer), and a virtual full mesh abstraction of the IP layer (on top, **2**). Since the cost function is based on link loads, these are first measured (**1**) on the actual network and then imported in the full mesh abstraction through the cost function (**3**). The (measured) traffic flows are then (re)routed (**4**) over the full mesh, and in a final step (**5**), we decide which IP links have to be set up (or even torn down). These decisions are signalled to the network, and routing information is updated for traffic flows that have been assigned a more optimal route. The route calculations are performed continuously, hence the term 'proactive', and eventually the routes (and therefore the logical topology) will settle down, assuming a correct choice was made for the cost function. The MTE process accepts dynamically changing traffic patterns (i.e. changing IP traffic flows). For example, removing some of the offered bandwidth will cause some IP links to drop in load, causing the MTE routing process to reoptimise the routing of the flows, and thus the logical topology.

Two more remarks need to be made concerning the loads that are used in calculating the costs. Firstly, links that are not currently available are assigned a high penalty cost (although they do not 'carry' any actual load). So the routing cost for a link is either dictated by the load in the actual IP logical topology, or by a penalty when the IP link is not available.

Secondly, such a load dependant cost function works fine on one condition: the load we use in calculating a route for a certain traffic flow (connection, LSP,…) is not the actual load on the IP links, but instead the load the IP links would have if it had to carry the traffic flow in question (note that both these loads are equal only for IP links that currently belong to the route of the traffic flow in the actual IP network). This can be illustrated with a simple example: suppose we have a certain traffic flow to reroute with a bandwidth of 50% of a lightpath (or, of an IP link). Assume there are two alternative IP links to route this flow over, one **(a)** with a current load of 0%, a second **(b)** with a current load of 40%. Neither currently carry the traffic flow. If we were to use the actual load as measured we would have:

for (a): load = 0%, thus a high cost (high cost penalty associated with low load)
for (b): load = 40%, thus a low cost (fairly moderate load)

This would lead us to believe that routing the 50% flow over link **(b)** is acceptable, however this overloads link **(b)** to 90% of its capacity (which already yields unacceptable QoS degradation in an IP network). Instead, keeping in mind the bandwidth of the traffic flow we are routing (=50%), we get:

for (a): load = 0%+50%, thus a low cost
for (b): load = 40%+50%, thus a high cost (correctly identifying the overload)

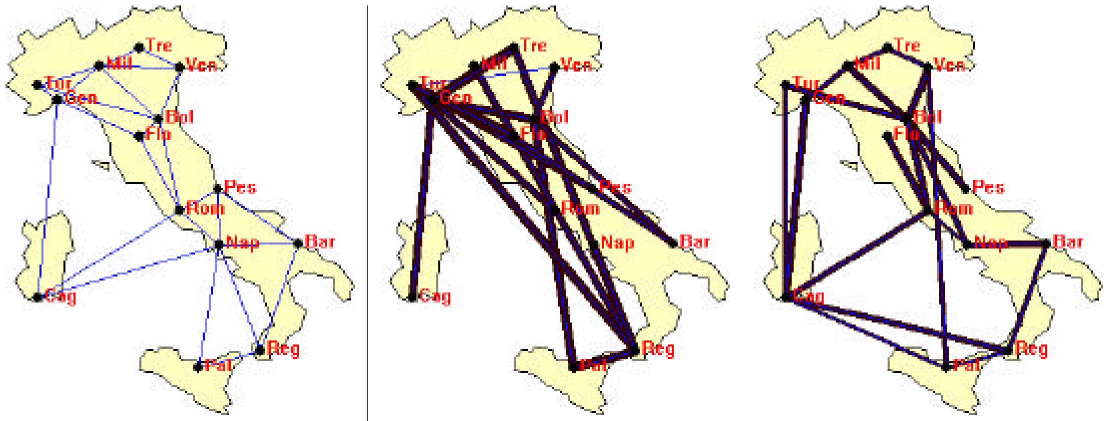
This way (basically, by using the anticipated cost of our routing of the flow over the network), we choose the appropriate IP link. This mechanism overcomes the apparent contradiction of assigning a high cost to lowly loaded IP links: the cost will only remain high (i.e. load = low) when trying to groom an insufficient number of IP traffic flows on an IP link. The effect of this can be seen on the load histogram of Fig. 1 (left); most IP links have a moderate load (meaning a fair number of flows have been multiplexed on them), while no IP links are used inefficiently or overloaded. This load distribution can be tailored using the **LLT**, **HLT** and **LMR** values. More information concerning this can be found in [5].

3. OPTICAL COST METRICS

The cost function and also the MTE strategy itself as described in the previous section do not require any knowledge of the optical layer, only IP layer link state information (load) is used. Therefore, all lightpaths are considered equally expensive to set up. In other words, it is assumed all lightpaths will cause a comparable load in the optical layer.

One may wonder whether it's not possible to incorporate some kind of optical metric into the MTE process, just as was done for the IP layer (cost function dependant on traffic load). This would allow spreading the lightpath load over the optical layer from within the MTE process, by appropriately choosing the lightpaths, since in most cases, there is some choice in setting up the IP logical topology for a given traffic demand. While most of these choices will yield about the same number of IP-links, some will be clearly inferior because the end-to-end lightpaths are much longer (geographically, and in number of hops).

In Fig. 2 we see what happens when we introduce an optical metric into the MTE process, running on a physical ASON topology detailed on the left part of the figure. The right part of the figure shows two logical IP topologies. Both route the same amount of traffic (IP load is signified by the thickness of the IP links): the left logical topology is the result of a MTE cycle ignoring optical layer metrics, while the right one uses a simple metric based on optical hop count (see below). While one may count a similar amount of IP links in both cases, once we plot these topologies with the nodes laid out geographically, the right one is clearly more 'open'; it resembles the physical topology and seems to consists of less long lightpaths (both in hop count and in a geographical sense).



No optical metric

Optical hop metric

Physical ASON topology

Logical IP topologies

Figure 2. Physical topology and sample logical topologies

The MTE process, which runs in the IP layer, does not see optical links, only node pairs between which there can be set up lightpaths. Thus, we calculate an optical cost metric for each node pair, and communicate this with the IP layer. Optical layer resource usage can be optimised using a *per lightpath* optical metric that rises with increasing optical hop count, such as the four sample linear optical metrics shown in Fig. 3. The optical hop count is determined by using shortest path routing in the optical layer. The simple linear cost metrics are of the form:

$$C + S.n, \qquad n \text{ the number of hops} \tag{2}$$

We have chosen a number of metrics such that the cost for a single hop lightpath is always the same (1 in this case). The constant C determines the fixed cost of a lightpath (regardless of number of hops). Fig. 3 shows C (and therefore S) for the selected optical metrics. For $C = 1$ (flat or no metric), there is only a fixed cost, and no additional cost for using hops in a lightpath. For $1 > C > 0$, the fixed cost dominates in the optical metric, while for $C < 0$, the hop count takes over. These optical metrics are never negative, since the minimum number of optical hops in a lightpath is 1, obviously, as shown on the right part of the figure.

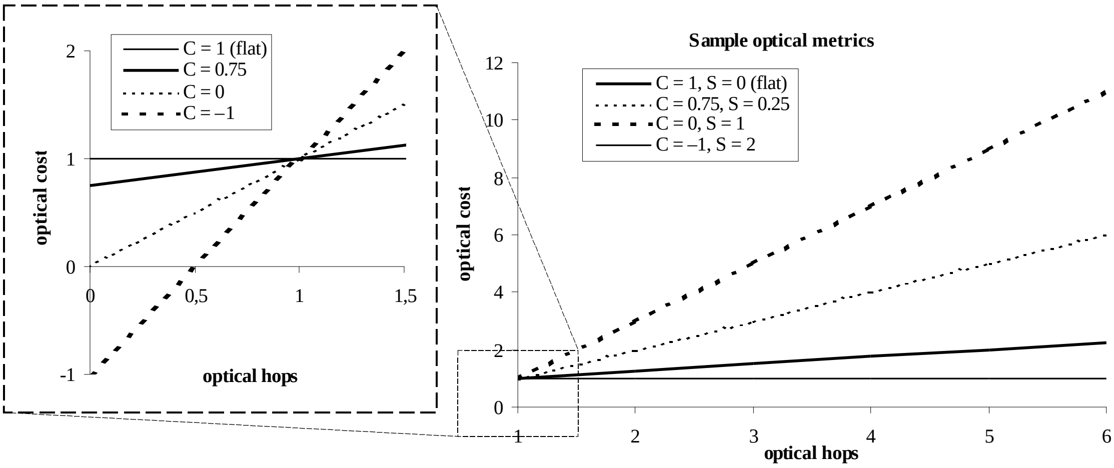


Figure 3. Sample optical metrics

Each IP link in the fictive full mesh used in the MTE strategy corresponds to a node pair in the optical layer. The optical cost metrics enable us to report a cost for each of these node pairs. In order to incorporate the optical metric, the MTE process now uses the product of both the IP cost function from section 2 and this optical metric as routing cost (instead of just the IP cost). Note that a flat metric corresponds with the case where we didn't consider the optical layer (as in section 2).

4. SIMULATION STUDY

If we now simulate the MTE process for a number of traffic sequences, we can see the influence of the optical metric on the optical layer resource usage. All of the consecutive traffic sequences consist of a number of traffic flows (with associated bandwidth demands $d_{i,j}$ for a flow between node i, j) to be routed over the logical IP topology. We designate a parameter $t_{n=1\ldots\mathrm{M}}$ corresponding with M discrete time intervals where these sequences are valid. The bandwidth of a demand $d_{i,\,j}$ (for a flow between node i, j) is distributed uniformly, with an upper limit $d(t_n)$:

$$\begin{aligned} d(t_n) &= rand[5\%...50\%] && \text{upper limit for time interval } n \\ d_{i,j}(t_n) &= rand[5\%...d(t_n)] && \text{traffic between node } i,\ j \text{ for } t_n \end{aligned} \tag{3}$$

Flow bandwidths are expressed as % of a lightpath, since this is the granularity with which the MTE requests IP link bandwidth – IP loads for the cost function are then in % of a lightpath (an IP link) as well. Assuming a full mesh of traffic flows, we can rewrite the traffic pattern (a collection of IP flows) as a traffic matrix $\mathbf{T}_n$ in equation 4, N the number of nodes in the IP layer.

$$\mathbf{T}_n = \begin{bmatrix} 0 & d_{1,2}(t_n) & d_{1,3}(t_n) & \cdots & d_{1,\mathrm{N}}(t_n) \\ d_{2,1}(t_n) & 0 & d_{2,3}(t_n) & \cdots & d_{2,\mathrm{N}}(t_n) \\ d_{3,1}(t_n) & d_{3,2}(t_n) & 0 & \cdots & d_{3,\mathrm{N}}(t_n) \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ d_{\mathrm{N},1}(t_n) & d_{\mathrm{N},2}(t_n) & d_{\mathrm{N},3}(t_n) & \cdots & 0 \end{bmatrix} \tag{4}$$

During the simulation, traffic patterns $\mathbf{T}_n$, n=1…M are applied in succession, meaning that with the introduction of a new pattern, the MTE strategy has to reconfigure the logical topology starting from what has been set up for the previous pattern. This leads to results more closely resembling normal operation, compared to optimising the network from scratch for each pattern. The MTE strategy, once traffic is applied, will converge to a logical topology with a suitable routing of the flows [5], for each time interval (and traffic pattern). Consecutive traffic patterns will have both a varying distribution of traffic flow bandwidths, but also a varying $d(t_n) \in [5\%\ldots50\%]$, so there can be large fluctuations in total IP volume to route from one stable logical topology to another – and thus a lot of lightpath set up/tear down. On the figures below, the results shown were achieved by averaging the performance of the MTE strategy for a large number of traffic patterns with a same total IP traffic volume characterised by $d(t_n)$. As such, we have called $d(t_n)$ the 'Max IP load' parameter on the X-axis, it is the maximum IP load a single traffic flow can cause on an IP link.

Fig. 4 shows the effect of the optical cost metric on the optical layer resource usage (left on the figure). We can clearly see a large improvement (especially for lower total volumes). Most of the optimisation happens when introducing the optical metric. Some further decrease in resource usage is seen when increasing the slope (raising the cost of long lightpaths). The improvement tends to diminish for higher traffic loads, because there traffic patterns are such that the MTE strategy requires an almost full mesh anyway, meaning our range of possibilities in choosing appropriate lightpaths is far less. The right part shows the optical cost savings, i.e. the decrease in number of point-to-point wavelength channels ('lambdas') due to introducing a optical cost metric (with C = 0, S = 1). Max #Lambda shows the decrease in maximum number of wavelength channels used simultaneously on any fibre, this stays fairly constant for varying total IP volumes. Total #Lambda indicate the total decrease in channels. We see a steady rise in total savings, up to about a 'Max IP load' of 30%. Beyond this point, the savings start to drop again, as the logical topologies now again start to become more highly meshed. Nevertheless, relative optical savings are quite significant, from 40% for lower IP volumes down to about 10% for higher volumes.

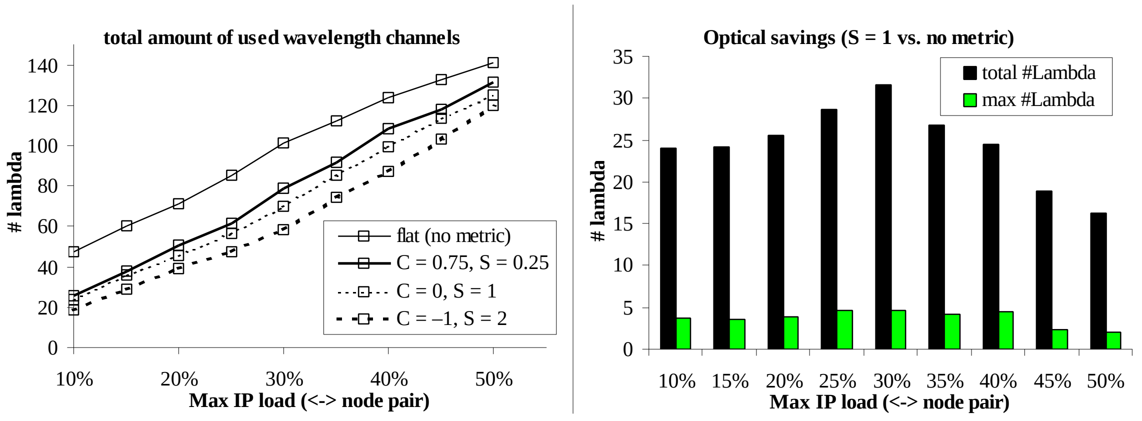


Figure 4. Optical layer performance

In all this, the total number of IP links in the logical topology, and thus the number of established lightpaths does not change when going from a flat metric to higher slopes. The optical savings are offset by a rise in IP router load however, as seen on Fig. 5 (left). As our cost metric shifts from a more fixed cost (C = 1) towards a cost dominated by the length of the lightpaths (C < 0), optimisation shifts more towards the optical layer, causing higher grooming in the IP layer. Note that IP router load is higher for lower IP volumes, as the MTE strategy there tends towards spanning tree topologies and full IP layer grooming.

We see a sudden increase once C drops below 0. This is confirmed when comparing the overall performance of IP and optical layer (right part of the figure). We see a gradual shift towards better optical layer performance for decreasing C, at the cost of slight increases in average IP router load, though once C becomes negative, the IP router load shoots up. A negative C corresponds with a fixed cost *bonus* for settings up lightpaths, leading to logical topologies consisting of a multitude of (very short) lightpaths, and more IP layer grooming. Adjusting C between 1 and 0 however, allows an operator to find a compromise between IP and optical layer performance and resource usage.

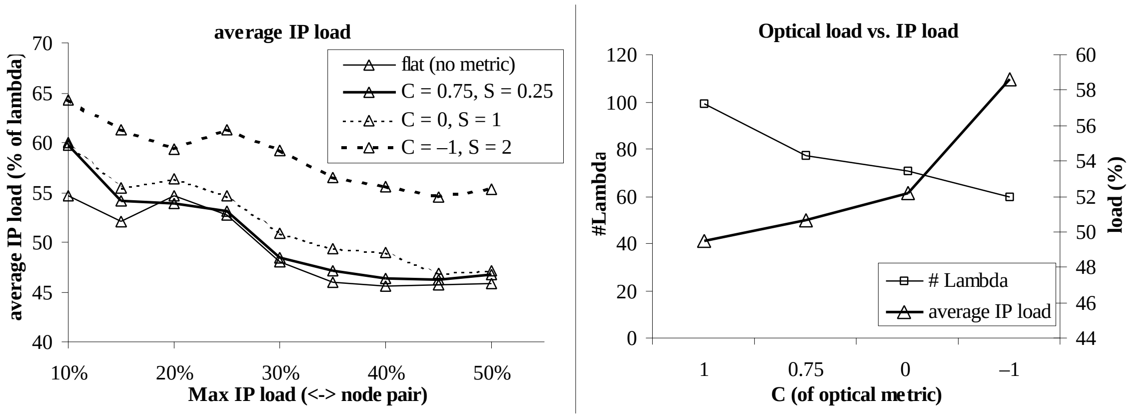


Figure 5. IP layer impact of the optical cost metric

5. CONCLUSIONS

We have explained how a MTE strategy, which for the most part uses IP link state information, can be extended to optimise optical layer performance as well, by using simple optical metrics. The simulation study in section 4 shows that optical layer performance can be improved significantly. Adjusting the optical metric reported to the IP layer allows one to compromise between IP and optical layer performance and resource usage.

REFERENCES

[1] D. Bishop et al.: The Lucent LambdaRouter: MEMS Technology of the Future Here Today, *IEEE Communications Magazine, vol. 40, nr 3, pp. 75-79, March 2002.*

[2] E. Mannie et al.: Generalized Multi-protocol Label Switching (GMPLS) Architecture, *Internet Draft, Network Working Group, 2002.*

[3] ITU-T Rec. G.8080: Architecture for the automatically switched optical network (ASON), *November 2001.*

[4] T. Cinkler et al.: Fairness Issues of Routing with Grooming and Shared Protection, *in Proc. ONDM 2004 pp. 665-684, Gent, Belgium, February 2004.*

[5] B. Puype et al.: Multi-layer Traffic Engineering in Data-centric Optical Networks, Illustration of Concepts and Benefits, *in Proc. ONDM 2003 – COST266/IST Optimist workshop pp. 211-226, Budapest, Hungary, February 2003.*

[6] IST Integrated Project Next generation Optical network for Broadband in Europe (NOBEL) http://dbs.cordis.lu/fep-cgi/srchidadb?ACTION=D&CALLER=PROJ_IST&QF_EP_RPG=506760

[7] IWT project Optical Networking and Node Architectures http://www.ibcn.intec.ugent.be/projects/IWT/GBOU_OPTICAL/index.html