

Comparison of IP/WDM transport network architectures for dynamic data traffic

(Invited)

M. Köhn¹, S. Bodamer², C.M. Gauger¹, S. Gunreben¹, G. Hu¹, D. Sass¹

¹Institute of Communication Networks and Computer Engineering, Univ. Stuttgart, Germany
Tel: +49 711 685-67991, Fax: +49 711 685-67983, E-mail: koehn@ikr.uni-stuttgart.de

²Ericsson GmbH, Backnang, Germany
Tel: +49 7191 133320, Fax: +49 7191 1363320, E-mail: stefan.bodamer@ericsson.com

In this paper, we outline key IP/WDM transport network architectures and present a framework for their comparison regarding resource requirements under dynamic data traffic. For enabling our dimensioning process by taking statistical multiplexing into account, we apply effective bandwidth theory to translate traffic demands with highly variable traffic characteristics into the required transport capacity. Finally, we use this framework to analyze the node and network resources for a German reference backbone network scenario with realistic traffic dynamics from traffic measurements.

1. Introduction

Solutions for next-generation IP backbone networks are a key issue for many network operators since IP will be the dominating network layer technology on which an ever-increasing number of applications with growing bandwidth requirements will be based. Without new network solutions, this trend would lead to a strong increase in number and size of IP routers while already today's requirements make it difficult to realize large-scale IP backbone networks in a stable and cost-efficient way. Also, those networks are increasingly leveraging on MPLS technologies (Multi-Protocol Label Switching) and thus the use of "IP network" in this paper in addition refers to combined IP/MPLS networks.

There are many benefits that an appropriate transport network technology can bring to future IP backbone networks. Beyond the fact that such a transport network will allow the operator to carry other services in parallel to IP services on a common transport platform, also a switched transport layer offers additional functionalities such as advanced restoration mechanisms or efficient bandwidth provisioning. Further, the capital expenditures (CapEx) for the IP network can be reduced by off-loading transit traffic from the IP layer to the transport domain as switching in the transport domain is significantly cheaper than packet forwarding on the IP layer. Finally, a managed transport layer underneath the IP/MPLS layer includes the potential for reduction of operational expenditures (OpEx).

In this paper we compare different network architectures with respect to the resources required on different layers to carry a given set of traffic demands. Several papers compare such IP/WDM network architectures assuming the given traffic demands to be identical on all layers and neglecting the impact of the dynamics of IP traffic [1][2]. Here, we consider the variability of data traffic. This strongly impacts the dimensioning process since it is

necessary to consider statistical multiplexing of different traffic streams on a link by appropriately dimensioning each link independent for to the carried traffic.

The remainder of this paper is structured as follows. In Section 2, we first provide a description of network architectures for IP/WDM transport networks. Further, we introduce in Section 3 a unified dimensioning process for multi-layer networks as well as introduce the calculation of the required capacity based on the effective bandwidth methodology for links that carry traffic with variable bit rate in Section 4. Then, in Section 5 we compare the different network architectures based on results of a numerical case study. Finally, Section 6 concludes the paper.

2. Transport Network Architectures

In this section we present different approaches that can be used to realize an IP/MPLS backbone network:

- **IP over WDM:** IP/MPLS routers are interconnected with point-to-point links.
- **IP over OCh:** IP/MPLS routers are interconnected by static optical channels (OCh, also called lightpaths) that traverse multiple hops.
- **IP over SDH/OTH:** A fully capable SDH/SONET/OTH transport plane is introduced under the IP/MPLS layer. Inter-connection on the optical layer is provided either by point-to-point WDM links („IP over SDH over WDM“) or by optical channels („IP over SDH over OCh“).

Figure 1 to 4 show these scenarios in a schematic way.

In Figure 1 the IP over WDM scenario is presented, basically consisting of IP/MPLS routers that are statically interconnected via point-to-point links usually based on WDM. For the interconnection between IP routers and the transmission systems, several technologies can be used. Most often, these links are either based on PoS (Packet over SONET/SDH) or on Gigabit Ethernet technology. For simplification purposes, in the context of this paper we assume a single IP router per location, while in reality there is usually a sub-network of routers present in every location due to reliance requirements and router port limitations.

The reference architecture for the “IP over OCh” scenario is shown in Figure 2. In each location wavelengths can be optically bypassed. The easiest way of realising this option is to use an optical patch panels (OPP). Other options include the usage of flexible all-optical cross-connects (OXC) or remotely configurable optical add-drop multiplexers (R-OADM). Optical bypass allows to set up optical channels/lightpaths between any node pair. Since these lightpaths are usually fat pipes with transmission rates of 10 or even 40 Gbit/s, connecting every node pair in this way is neither required nor cost-efficient for a reasonable amount of traffic. Therefore intermediate grooming of selected traffic flows in the electrical domain is required which is done in the IP/MPLS router in the IP over OCh scenario. Transit traffic is thus not completely removed from the IP/MPLS layer.

In the IP over SDH scenarios (Figure 3 and Figure 4) a flexible transport network is introduced underneath the IP/MPLS layer. This layer is based on electronically switched TDM circuits, available with technologies such as SDH, SONET, or OTN. In an opaque network layout (Figure 3) these electrical cross-connects (EXCs) are interconnected by WDM systems as described above. Alternatively the EXCs have long reach “colored” line cards, allowing a direct access to the WDM system without the need for short reach “grey” interfaces and separate transponder cards in the WDM system. Compared to the scenarios above, the full transport scenario enables setting up direct SDH/SONET/OTN connections between IP routers. Since these pipes can have sub-wavelength granularity (e. g. STM-1 or STM-16) it is possible to build up a full mesh virtual topology thereby avoiding any transit

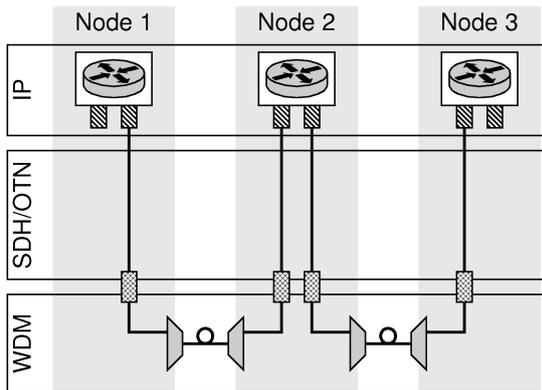


Figure 1: IP over WDM

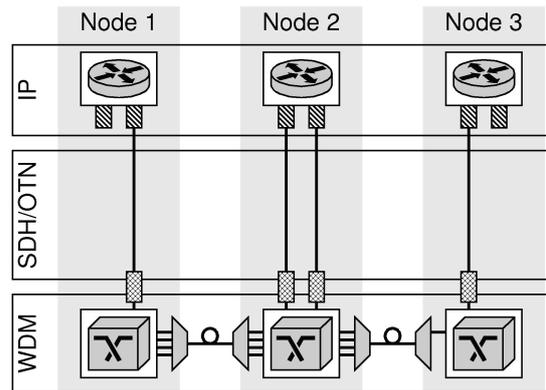


Figure 2: IP over OCh

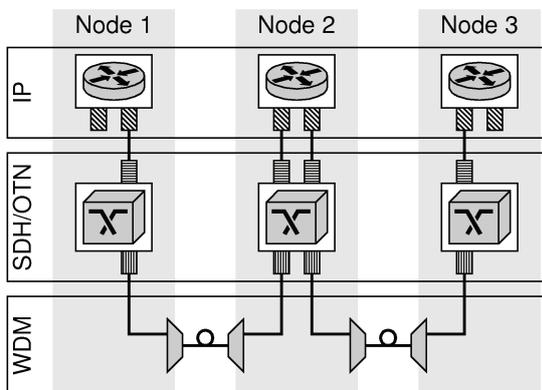


Figure 3: IP over SDH over WDM

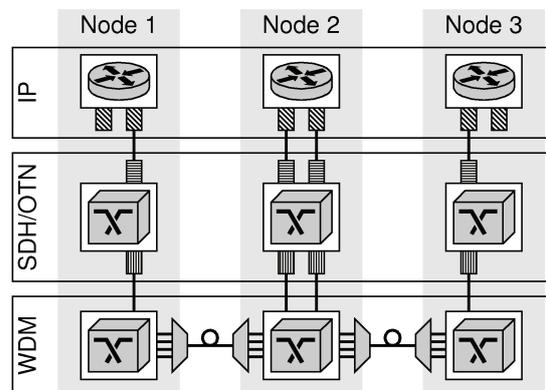


Figure 4: IP over SDH over OCh

traffic in the IP layer. Mapping of data flows into transport network connections may be done in the IP routers or in the EXCs using data cards that are capable of mapping L2 flows into SDH Virtual Containers (VCs).

While in the opaque scenario shown in Figure 3 EXCs are interconnected by point-to-point WDM links, optical bypass can also help in the IP over SDH case to reduce network cost (Figure 4). In this so-called IP over SDH over OCh scenario, each node has an additional capability to transparently bypass a wavelength channel in the optical plane without terminating it at the EXC. However, all channels can be terminated if required by the traffic matrix or if beneficial for efficient grooming into wavelength channels.

3. Unified Dimensioning Process

In this section we introduce a unified dimensioning process for layered networks. This process derives the resources, which are required to carry a given traffic demand with a given QoS objective, for each layer network. We will use this process in Section 5 to quantitatively compare the different architectures described in Section 2.

The dimensioning of a layer network accounts for node resources, e.g., switching capacity, as well as link resources, e.g., number of interfaces and transmission capacity. In multi-layer network architectures, this transmission capacity is again provisioned by the underlying layer network. Only in the lowest layer network, it directly matches a physical transmission channel, e.g. a wavelength channel. In techno-economic studies, the cost of physical transmission is commonly neglected as it is dominated by the cost of node resources and interfaces [3].

Figure 5 depicts a basic dimensioning process for layer network n in a multi-layer network. We dimension each layer network starting from demands as specified by the user traffic demands or the link resources of the adjacent layer network $n+1$. In a flow analysis, we route these demands on the link topology of the layer network n according to its routing or grooming scheme to obtain the link loads. Then, we compute the resources from the link loads for a given QoS objective. These resources then define the demands for layer network $n-1$. Finally, we derive the required switching capacity as the sum of attached link capacities for each node and the required number of interfaces.

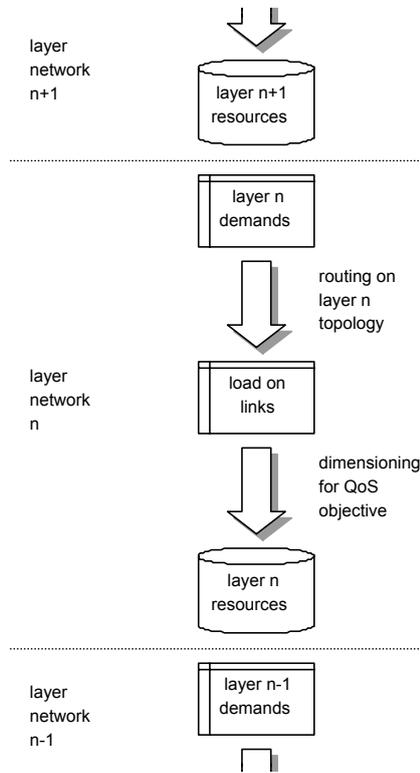


Figure 5: Unified dimensioning of multi-layer networks

We formulate this network dimensioning process as a sequence of matrix transformations for routing and resource computation. Routing mostly depends on the applied routing scheme, e.g., shortest path routing. In contrast, the resource computation does not only depend on the QoS objective but also on the traffic characteristics and the network technology. For static or pseudo-static layer networks no resource computation is necessary as the link capacities are identical to the link loads. For dynamic layer networks, we assume independence of links and a homogeneous case, i.e., identical traffic characteristics and QoS objectives on all links. In this case, we can apply the same traffic model to all links independently, which greatly simplifies resource computation. The following section will describe a well-known model for resource computation in packet-switched networks using effective bandwidth.

The dimensioning process outlined so far may not be resource optimal and practical network implementation may require further calculations. In order to avoid underutilized links, an additional optimization step can reroute demands traversing such links. In otherwise strictly single-hop networks, this may require to allow multi-hop paths, a concept commonly referred to as grooming. Similarly, the fixed capacity of transport circuits, e.g., STM connections or lightpaths, requires discretization of the resources. Beyond mathematical rounding, advanced algorithms can globally minimize the difference of discretized and original results [4]. In the process outlined above, both can be modeled by additional transformations after the routing or the resource computation steps, respectively.

4. Effective Bandwidth Modeling

Dynamic data traffic enables the possibility to achieve a statistical multiplexing gain by aggregating traffic, leading to smaller overall bandwidth requirements. The required bandwidth demands can be estimated by the so-called *equivalent capacity*, incorporating the QoS constraints (loss, delay) and the traffic source statistical characteristics by means of effective bandwidth¹.

¹ According to [5], we use the term effective bandwidth as a pure property of a traffic source whereas equivalent capacity applies in a system context reflecting both properties of source and system.

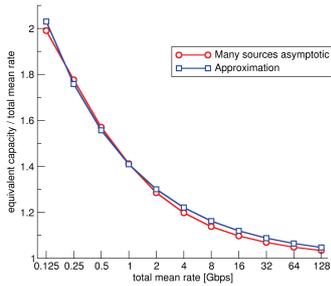


Figure 6: dimensioning curve

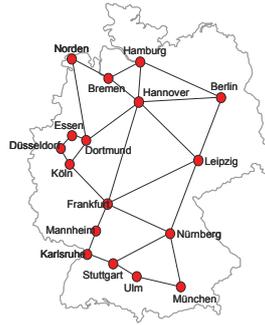


Figure 7: network topology

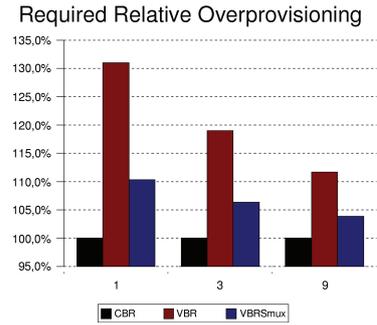


Figure 8: impact of traffic

A key element for the calculation of the equivalent capacity is the effective bandwidth. In general, the effective bandwidth is a function $\alpha(s, t)$, with space parameter s and time parameter t , describing the statistical properties of the traffic source [5]. Several effective bandwidth schemes exist, ranging from rather general, complex expressions [5] to schemes assuming certain traffic models [6], [7].

These schemes can be applied to the Many Sources Asymptotic (MSA), representing a powerful method for determining the equivalent capacity [8], [9], where the traffic dynamics of the sources are described by the corresponding effective bandwidth functions. The idea behind MSA is that the “system” is scaled by a size parameter and it is considered as asymptotically large. This leads to an optimization formula, where the extreme point is reached in the operating point (s^*, t^*) . With this, the equivalent capacity C^* can be derived based on $\alpha(s^*, t^*)$.

An important property in this context is that for small variations of the traffic stream mix, no major variation of the operating point happens, i.e., the operating point remains stable with respect to small variations of the traffic mix [8].

With this concept, it is possible to dimension links for any mean rate m by calculating the equivalent capacity based on the effective bandwidth of the source. Nevertheless, this is only possible, if the used reference source can be directly scaled by its mean rate m . But this is not always the case as, e.g., for traces where the mean rate can be inferred, however not changed.

In this case, links can be dimensioned using a reference source for which the traffic dynamics are given by means of its effective bandwidth. Then, the equivalent capacity can be calculated based on a scaled system, i.e., theoretical superposition of this reference source by means of the linearity property of the effective bandwidth. From this scaling, now, we are able to map any given mean rate m to the appropriate equivalent capacity C^* based on the traffic dynamics given by the reference source.

In Figure 6, the ratio of equivalent capacity over total mean, i.e., the over-provisioning factor is plotted over total mean m . For the red curve, we used as the reference source for the traffic dynamics measurement data from an uplink of a dormitory network of about 1000 users at the University of Stuttgart, where the users have already access link rates of 100Mbps [10]. The equivalent capacities are calculated on basis of MSA. This curve can be well fitted by the function $\tilde{C}^*(m)/m = 1 + k_1 m^{k_2}$ with appropriate constants k_1, k_2 . This is represented by the blue curve in Figure 6 and is further on used in this paper due to complexity and numerical reasons.

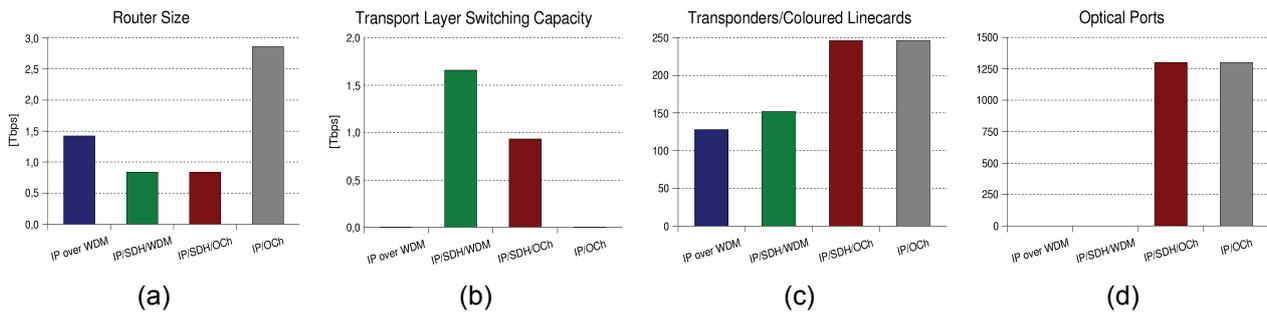


Figure 9: Results for scenario without optimization

5. Numerical Case Studies

In this section we compare the different network architectures introduced in Section 2 by means of node capacities and resources. We apply the network dimensioning procedures described in Section 3 for dynamic data traffic according to Section 4. We furthermore use for our study a realistic, but hypothetical reference network, representing a German national backbone network [11]. This network scenario comprises 17 backbone node locations, which are inter-connected with a meshed topology shown in Figure 7.

In the following, we first evaluate the impact of the traffic dynamics with and without exploiting the statistical multiplexing gain. In a second step the capacities and resources are compared for the different network architectures with multiplexed traffic demands.

5.1 Impact of traffic characteristic

Figure 8 depicts the required relative over-provisioning of IP router capacity for the IP over WDM architecture for three scenarios CBR, VBR and VBRSmux, respectively and three different load situations. The scenarios can be characterized as follows:

First, in the CBR scenario we understand rates given in the traffic matrix as the end-to-end demands with constant bit rate. So, neither dynamics nor statistical multiplexing must be considered. Second, in the VBR scenario we understand the rates as mean of data streams with variable bit rate. We consider the traffic dynamics in the dimensioning process but neglect the benefits from statistical multiplexing. Third, in order to show statistical multiplexing gain in the network in the VBRSmux we again understand scenario the rates as mean of a variable bit rate stream and statistically multiplex the individual streams sharing a link.

In order to show the impact of the overall load in the network, we scaled the traffic demand matrix linearly by a factor of 3 and 9, respectively. All results are normalized to the required provisioning of the CBR scenario.

For the first load scenario, we find that the network has to be over-dimensioned by about 30% for the VBR case and by about 10% for the VBRSmux case. So, in both VBR scenarios, the required over-provisioning is higher than for the CBR case. This is due to the fact that the CBR scenario is equivalent to a VBR scenario where the aggregation smoothes the dynamic behavior and thus can be seen as limiting case. Also, comparing the VBR scenarios, it can be seen that taking statistical multiplexing into account the required over-provisioning is significantly reduced.

In the remaining load situations the relative over-provisioning is reduced in case of VBR and VBRSmux traffic due to the decreasing variance by increasing load.

In the following, we use the 100% load scenario and consider only the VBRSmux traffic scenario.

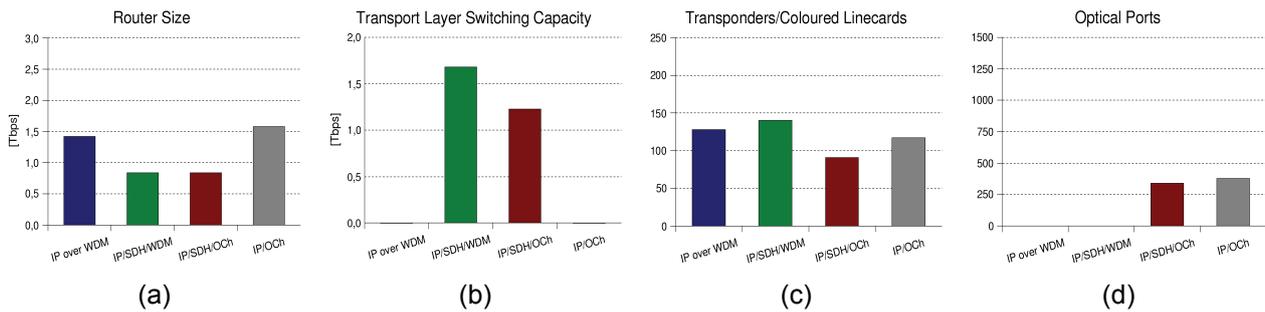


Figure 10: Results for scenario with optimization

5.2 Network dimensioning

In this section, we compare the different network architectures quantitatively without optimization. In Figure 9, we show for the different architectures the sum of the size of all routers (a), the sum of transport layer switching capacity (b), the total number of transponders/colored line cards (c) and the total number of optical ports of an OXC (d).

Comparing the router sizes in Figure 9a, it can be seen that the total capacity required in the IP over WDM scenario is about 1.4 Tbps. Here, the virtual topology in the IP layer matches the physical topology by point-to-point links of neighbor nodes and routing is only performed on the IP layer. In the IP/SDH/WDM and IP/SDH/OCh scenarios, the capacity is reduced to 0.8 Tbps as additional switching is provided in the transport layers and transit traffic is off-loaded to these layers. The fourth bar shows the amount of router capacity for the IP/OCh architecture, which is three times larger than in the IP only case. That is because IP/OCh establishes a virtual full mesh on the IP layer and for each connection an IP line card has to be installed.

Figure 9b shows the needed transport layer switching capacity. As the IP over WDM and the IP/OCh scenario perform no switching on this layer their bars are missing. Comparing the IP/SDH/OCh to the IP/SDH/WDM architecture, the capacity is reduced by 0.5 Tbps as additionally switching is performed in the optical layer off-loading through traffic from the electrical layer.

In Figure 9c, it can be seen that the number of transponders is minimized, when the virtual topology matches the physical topology. The nodes are connected by point-to-point links, which in total results in about 125 transponders. When introducing a grooming layer, i.e., the IP/SDH/WDM scenario, the number of transponders is slightly increased which results from a lower traffic aggregation on each channel. Finally, in the IP/OCh and IP/SDH/OCh scenarios the maximum number of transponders is needed. Here, the optical layer is forming a (almost) full mesh topology.

The size of the optical cross connects in the IP/OCh and the IP/SDH/OCh scenario is shown in the remaining bar chart (Figure 9d). Both scenarios do not differ in the number of optical ports, because of the full mesh topology.

5.3 Network dimensioning with optimal grooming

In Figure 10, the results with optimal grooming are depicted for the same scenarios. This grooming scheme tries to minimize the number of transponders while keeping the router and switch sizes low [1].

Comparing the results with and without optimal grooming, the results for IP over WDM and IP/SDH/WDM are almost unchanged. In the latter scenario, only the number of colored line cards is slightly decreased.

In contrast, in the IP/OCh and IP/SDH/OCh scenario a reduction of the total number of optical ports and the number of line cards can be observed by a factor of five and two,

respectively. The reason is that due to the optimization in the optical layer a full mesh topology is not necessary anymore. The price for this is the increase of electrical layer switching capacity in the IP/SDH/OCh scenario as now traffic must be switched in the EXCs, which has been transmitted on a direct optical channel in case of no optimization.

Further, in the first graph it can be seen that the IP router size in the IP/OCh is nearly halved compared to the case without optimization. This is due to the fact that the router sizes are calculated based on their port capacity – tributary and trunk ports – and not based on the processed traffic. As mentioned, in the IP/OCh architecture the grooming algorithm reduces the number of transponders and thus also the number of line cards in the IP routers. Nevertheless, in this study the routers are still larger than in the IP over WDM scenario. This is due to the optimization algorithm, which was developed for scenarios without considering statistical multiplexing.

6. Conclusions

In this paper, we have described four reference network architectures for IP transport. Further, we have explained a unified dimensioning process for multi-layer networks that considers the traffic dynamics based on the effective bandwidth theory.

In a numerical evaluation, we have shown the impact of traffic dynamics for dimensioning as well as the impact of the network architecture on the resource requirements in all modeled layers based on the characteristics taken from real traffic measurements. Concluding, the introduction of a transport network can reduce the size of the routers by off-loading the transit traffic at the cost of additional equipment required in the transport layer.

Acknowledgements

The work described in this paper was funded by the EU IST Integrated Project NOBEL (FP6-506760). The authors would like to thank Aurelién Galicher for his work.

References

- [1] J. Späth, S. Bodamer, C. Glingener: "Next-generation transport solutions for IP backbone networks - benefits of an ASTN-based multi-layer OTN network", *Proc. of Asia-Pacific Optical Communications (APOC)*, 2004.
- [2] R. Batchellor, O. Gerstel, "Cost Effective Architectures for Core Transport Networks", *Proc. of the 2006 Optical Fiber Communication Conference (OFC2006)*, PDP42, 2006.
- [3] S. Bodamer, J. Späth, C. Glingener, "An Efficient Method to Estimate Transponder Count in Multi-Layer Transport Networks", *Proc. of the IEEE Global Telecommunications Conference (Globecom)*, pp. 1780-1785, 2004.
- [4] M. Barry, S. Bodamer, S. Brink, L. Levrau, J. Späth, "Dynamic Multi-Layer Networks - Architecture, Modelling, Tool Support", Technical Report, Marconi GmbH, 2003
- [5] F. Kelly, "Notes on Effective Bandwidths," In: F. P. Kelly, S. Zachary, and I. B. Ziedins (eds.), *Stochastic Networks: Theory and Applications*, Royal Statistical Society Lecture Note Series, vol. 4, Oxford University Press, 1996, pp. 141-168.
- [6] R. Guérin, H. Ahmadi, "Equivalent Capacity and Its Applications to Bandwidth Allocation in High-Speed Networks", *IEEE J. Select. Areas Commun.*, vol. 9, no. 7, pp. 968-981, Sept. 1991.
- [7] S. Bodamer, J. Charzinski, "Evaluation of Effective Bandwidth Schemes for Self-Similar Traffic," *Proc. 13th ITC Specialist Seminar on IP Measurement, Modeling and Management*, Monterey, CA, September 2000, pp. 21-1-21-10.
- [8] C. Courcoubetis, R. Weber, "Buffer overflow asymptotics for a switch handling many traffic sources", *Journal of Probability*, 33, pp. 878-897, 1996.
- [9] D. Botvich, N. Duffield, "Large Deviation Theory, the shape of the loss curve, and economies of scale in large multiplexers", *Queueing Systems*, 20, pp. 293-320, 1995.
- [10] D. Sass, "Internet traces of the Selfnet university dormitory network, Trace UST2", University of Stuttgart, IKR, 2004, <http://www.ikr.uni-stuttgart.de/~sass/traces/>
- [11] A. Betker et al. "A Set of Typical Transport Network Scenarios for Network Modelling", *Proc. of the 5. ITG Symposium on Photonic Networks*, pp. 65-72, 2004