

Benchmarking and viability assessment of Optical Packet Switching for metro networks

C. Develder, A. Stavdas, A. Bianco, D. Careglio, H. Lønsethagen,

J. Fernández-Palacios, R. Van Caenegem, S. Sygletos, F. Neri, J. Solé-Pareta, M. Pickavet,

N. Le Sauze, P. Demeester

Abstract—Optical Packet Switching (OPS) has been proposed as a strong candidate for future metro networks. In this paper we assess the viability of an OPS-based ring architecture as proposed within the research project DAVID, funded by the European Commission through the IST-framework. We discuss its feasibility from a physical layer point of view, and explore its limitations in size. Through dimensioning studies, we show that the proposed OPS architecture is competitive with respect to alternative MAN

Manuscript received December 12, 2003. This work has been funded by the European Commission through the IST-Project DAVID (IST-1999-11387), and by the Flemish Government through the IWT GBOU-project “Optical Networking and Node Architectures”. C. Develder has been supported as a Research Assistant of the Fund for Scientific Research – Flanders (F.W.O.–VI.), Belgium.

C. Develder is with OPNET Technologies, 9000 Gent, Belgium (e-mail: cdevelder@opnet.com)

R. Van Caenegem, M. Pickavet and P. Demeester are with the Department of Information Technology (INTEC), Ghent University, 9000 Gent, Belgium (e-mail: {ruth.vancaenegem, mario.pickavet, piet.demeester}@intec.ugent.be).

A. Stavdas and S. Sygletos are with the Institute of Communication and Computer Systems ICCS/EPISSEY, National Technical University of Athens, Athens 157 73, Greece.

A. Bianco and F. Neri are with Dipartimento di Elettronica, Politecnico di Torino, Torino 10129, Italy (e-mail: {fabio.neri, andrea.bianco}@polito.it).

D. Careglio and J. Solé-Pareta are with with the Departament of Computer Architecture, Universitat Politècnica de Catalunya, Barcelona 08034, Spain (e-mail: {careglio,pareta}@ac.upc.es).

H. Lønsethagen is with Telenor, Norway.

J. Fernandez-Palacios with the Technology Strategy Department of Telefonica I+D, 28043 Madrid, Spain (e-mail: jpfpg@tid.es).

N. Le Sauze is with Alcatel Research and Innovation, Route de Nozay, 91460 Marcoussis, France.

approaches including SDH, RPR, and star-based Ethernet. Finally, we discuss the proposed OPS architectures from a logical performance point of view and explain a high quality scheduling algorithm to control the packet switching operations in the rings.

Index Terms—Optical packet switching, Wavelength division multiplexing, Metropolitan Area Networks, performance, MAC.

I. INTRODUCTION

DESPITE the recent economic malaise, the demand for telecommunication services continues to grow steadily. Even though this growth may have been over-enthusiastically acclaimed (leading to the creation and explosion of the “bubble”), it cannot be denied that telecommunication networks are at the heart of our information-based economy and society. These networks nowadays are largely based on optical fibre technology. Indeed, the use of Wavelength Division Multiplexing (WDM) offers massive bandwidth through the parallel transmission of high bitrate channels onto the same fibre, at a very attractive cost per bit. Currently, we are witnessing the shift from purely point-to-point WDM systems to the introduction of real networking functionality at the optical level.

The first steps in that direction are being taken with the development of Automatically Switched Optical Networks (ASON), enabling the automated set-up and tear-down of so-called lightpaths. Wavelengths are set-up between endpoints in the WDM network, avoiding costly E/O conversions in intermediate nodes. While this undoubtedly is a great step forward, the resulting network is still relatively static and mandates efficient aggregation and grooming techniques. Thus, it may be suitable for a core network—carrying highly aggregated traffic streams that are relatively predictable—but only to a far lesser extent in a Metropolitan Area Network (MAN) environment. A MAN needs to provide a large variety of service qualities in a highly dynamic environment where the cost per termination (rather than cost per bit) tends to be dominant. Traffic of dissimilar protocols and bitrates needs to be carried, fluctuating heavily in both volume and space. Traditional approaches such as SDH/SONET do not offer sufficient flexibility, and are optimized for voice circuits rather than the now dominant packet switched data traffic.

Optical packet switching (OPS) can offer the flexible and bandwidth efficient architecture that is called for, even though the actual deployment of OPS in future high performance networks is still questioned by some researchers [1]. Yet, compared to circuit switched approaches, it provides smaller granularity to the optical layer (on a packet-by-packet basis, thus allowing a high degree of statistical multiplexing), while still allowing for optical bypassing of transit nodes (without E/O conversions) for traffic traversing multiple hops. Ideally, we envisage transparent optical networking, where the optical packet can contain an arbitrary client layer protocol. For a recent overview of OPS architectures and their pros and cons, we refer to [2] and references

therein. OPS-based architectures for metropolitan environments are discussed below.

Various research projects have already proposed MAN architectures based on OPS, often proposing a pragmatic combination of electronics for controlling the switch, and optics for the actual switching. Multiple proposals have been described using ring-based networks comprising bufferless optical nodes because of the difficulty in implementing memory in the optical domain. In the HORNET architecture defined at Stanford [3], multiple wavelengths are deployed onto a single fibre ring (in WDM), where each node can receive on only a single (fixed) wavelength, being equipped with tunable transmitters and fixed receivers. Transit packets crossing a node at its receiver wavelength are dropped regardless of their final destination. This leads to a multi-hop scheme requiring excessive E/O conversions, electronic buffering and packet processing. Switching decisions are made based on packet headers which are sent using orthogonal FSK modulation.

A Japanese research project [4] demonstrates a slotted OPS architecture based on 2×2 switches to put packets on and off the ring, using a single wavelength for packet transmission, and an extra control wavelength carrying the associated headers. A central master node regulates access to the ring by generating empty slots marked with an address of the node that is granted permission to use it.

Also in Europe, several projects have proposed OPS for MAN. Among them, a Dutch project FLAMINGO [5] is quite similar to the Japanese project, as it is also based on 2×2 switches and a dedicated control channel wavelength. The Italian RINGO project [6] used a unidirectional slotted WDM/TDM architecture with fixed receivers and tunable transmitters: each node has its own dedicated wavelength for packet reception. Thus, there is no need for switching components: (de)multiplexers and passive coupling of light from a tunable transmitter suffices.

The work presented in this paper discusses the MAN architecture devised in the frame of the DAVID project [7,10]. Two alternative architectures for MAN rings have been proposed, and will be described in more detail in the following section. Any of those OPS approaches, when mature for commercial deployment, will naturally have to compete not only with SONET/SDH, but also with other recent MAN technologies such as Ethernet (IEEE 802.3) or Resilient Packet Rings (RPR, IEEE 802.17). Therefore, we will not restrict ourselves to detailing the DAVID architectures and their performance, but also include benchmarking studies comparing them against non-OPS technologies.

The remainder of this paper is organized as follows: the subsequent Section II will outline the DAVID architectures. Their feasibility from a physical performance point of view will be addressed in Section III. From a cost perspective, they will be benchmarked against the aforementioned alternatives in Section IV. A discussion of the logical performance will be presented in Section V. All conclusions will be summarized in the final Section VI.

II. THE DAVID NODE AND NETWORK ARCHITECTURES

The IST-project “Data And Voice Integration on DWDM” (DAVID) aims at proposing a viable approach towards OPS, by developing networking concepts and technologies for future optical networks [7]. The work ranges from theoretical studies covering traffic studies, control, scheduling algorithms, MAC protocols, etc. over studies of physical feasibility and advanced optical components, to a proof-of-concept demonstrator. The studied network scenarios also cover backbone scenarios, but much attention was given to the development of a MAN architecture, given the particular opportunities in such a context [8]. A generic view of the DAVID architecture is given in Fig. 1.

The DAVID MAN comprises multiple physical rings interconnected through a so-called Hub. A ring will comprise one or more fibres, each operated in DWDM regime (10 Gbit/s per channel has been assumed for the studies presented in this paper). One wavelength constitutes a dedicated control channel, while other wavelengths are used to carry the actual data in the form of fixed length packets. A time slotted operation is used, since synchronous network operation is considered easier to implement. The adopted ring architecture thus uses both wavelength division multiple access (WDMA) and time division multiple access (TDMA).

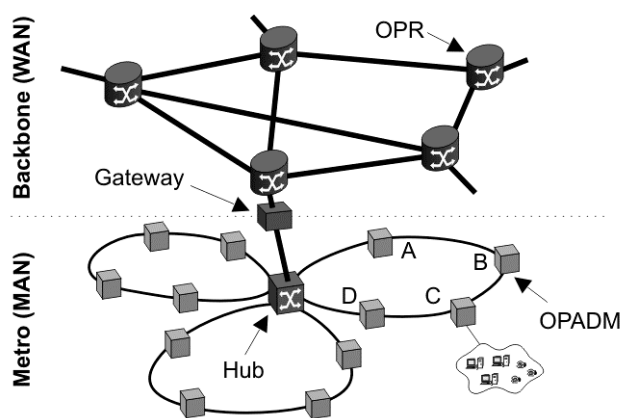


Fig. 1. Generic view of the DAVID network architecture.

The Optical Packet Add/Drop Multiplexer (OPADM) ring node puts optical packets [containing client layer traffic, e.g. IP] on the ring, using a Medium Access Control (MAC) protocol to decide which timeslot at what wavelength to use. By enforcing proper constraints via the MAC protocol (see further, Section V.B), contention on the optical packet level is avoided, and the need for buffering on the optical path within the MAN is eliminated: all buffering is done electronically in the add/drop interfaces. At the optical level, each node transceiver, even if tunable, is capable of transmitting and receiving on only one channel at a time (i.e., its bandwidth is equal to the capacity of one channel). Thus a good compromise is achieved between

optical and electronic technologies, keeping the high-speed electronic data path at an acceptable level of complexity.

The Hub, which also is bufferless, forms the interconnection point of multiple rings and provides access towards the WAN through a gateway. This WAN-connection from a logical point of view can be seen as an extra ring to switch traffic to and from. The gateway will be responsible of solving contention between packet flows between MAN and WAN. The latter consists of Optical Packet Routers (OPRs) interconnected in a meshed topology. In contrast to the MAN, an OPR in the WAN may exploit optical buffers in the form of Fibre Delay Lines (FDLs) to aid in contention resolution [9].

The Hub performs switching of the entire metro network's capacity, transferring packets from any incoming to any outgoing fiber, and performing wavelength conversion if required to achieve spectral adaptation/matching for a source/destination OPADM pair. In DAVID, the switching matrix in the Hub was based on the same architecture as an OPR in the WAN (but without the FDL buffers).

For the MAN ring nodes however, two alternative architectures have been studied in DAVID. The first is a *Passive* architecture, relying on commercially mature, low-cost technology, using only passive optical components. As sketched in Fig. 2, the architecture is extremely simple. At the OPADM's input, the control channel first is split off, followed by a fibre delay line (FDL) for the data channels to account for the delay in processing control information. This processing is achieved by O/E/O conversion of the control channel and implements the MAC decision protocol. On the data path, a simple 2:2 coupler is used (i) to add packets by coupling light coming from a burst mode transmitter, and (ii) to drop packets by guiding light to a burst mode receiver. The WDM channels from multiple receivers and transmitters are separated resp. combined by demultiplexers resp. multiplexers. To allow simultaneous add and drop operations within the same timeslot, up- and downstream traffic channels are spectrally separated [which in addition obviates crosstalk between add and drop channels]. Clearly, the Hub will need to perform wavelength conversion from the send to the receive spectrum to allow communication.

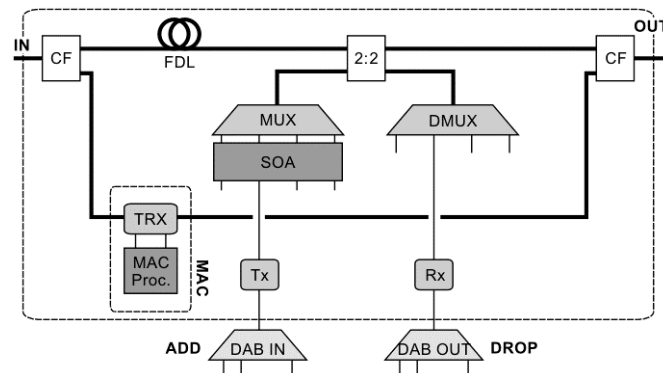


Fig. 2. The Passive OPADM node structure. (CF: control channel filter; TRX: MAC transceiver; Proc.: MAC processor; FDL: Fibre Delay Line; SOA: Semiconductor Optical Amplifier array; SOA+Tx: fixed wavelength burst mode transmitter; Rx: fixed wavelength burst mode receiver; DAB: Data Aggregation Board; DAB IN + DAB OUT: aggregation node).

The main drawback of the Passive architecture is that packets cannot be physically removed from the wavelength comb, and therefore propagate past their final destination (prohibiting reuse of the same slots for transmission). Consequently, the Hub needs to take care of packet erasure from the ring. In addition, so-called space reuse is impossible with a Passive structure: since all traffic necessarily needs to cross the Hub, the same slot cannot be reused for non-overlapping connections on the same ring (e.g. from A to B, and from C to D in Fig. 1). Moreover, spectral separation of up- and downstream traffic doubles the amount of required wavelengths.

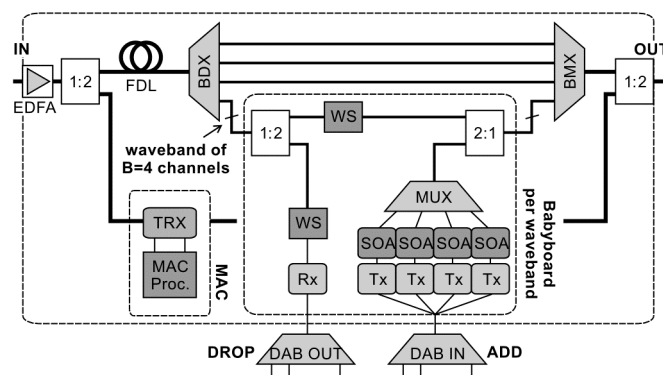


Fig. 3. The Active OPADM node structure. (1:2: 1x2 splitter; EDFA: Erbium-Doped Fibre Amplifier; FDL: Fibre Delay Line; BDX: waveband demultiplexer, BMX: waveband multiplexer splitting/combining disjoint groups of B wavelengths; WS: wavelength selector for B channels with pass/no-pass gates, here B=4; other acronyms: same meaning as in Fig. 2).

More advanced components are used in the *Active* node structure outlined in Fig. 3, which in addition employs a waveband concept. Instead of a passive coupler, we find a waveband demultiplexer (BDX), isolating light in groups of B wavelengths per band (unless otherwise stated, we assume B=4). For each waveband addressed in a ring node, a so-called babyboard is installed.

That board comprises a single receiver and a transmitter that is tunable over the B wavelengths in a particular waveband (in Fig. 3, this is implemented through an array of transmitters with semi-conductor optical amplifier (SOA) selectors to keep only one signal). It further permits selective erasure of packets by the gates in the wavelength selector (WS) block on the through path. Apparent advantages of this Active structure are that it allows for slot re-use (i.e. the dropped slot can be re-used for transmission of new data), it does not require separation of up- and downstream traffic, and a flexible use of the WDM domain (cf. waveband design and tunability of transceivers per waveband). The main drawback of the Active structure clearly is its higher initial cost, yet its modular structure based on the babyboard concept may allow for longer term savings because of the pay-as-you-grow approach. In the next section, we will study both Active and Passive structures from a physical layer point of view.

III. PHYSICAL LIMITATIONS

The current physical layer benchmarking study was carried out assuming a MAN consisting of four rings. All system and component parameters were scaled to support a total capacity of 1.28 Tb/s. Using a line rate of 10 Gbit/s per wavelength channel, this implies that each ring supports 32 channels. In order to obtain generally applicable conclusions, the physical layer performance should be highly independent of the traffic pattern. Therefore, two assumptions were made. First it was postulated that any OPADM has the potential to add/drop all the MAN capacity and so 32 transceivers per OPADM were used in the simulations. Secondly, the ring is operated at the maximum capacity at all times, where we assume the entire upstream capacity (32 channels) to be added in the first node. This assumption might be seen as the worst-case networking scenario, esp. for the Passive architecture with its absence of spatial slot re-use, since none of the slots are left over for adding traffic from nodes further upstream. Compared to [10], the physical layer simulations carried out in this work are based on components and system parameters that are either identical or very close to the ones measured in the corresponding sub-systems of the DAVID demonstrator [11].

The aim of this section is to identify important physical layer performance issues like the maximum transparent MAN ring length and the total number of optical nodes in a cascade. This is done using optical components with realistic performance e.g. Erbium Doped Fibre Amplifiers (EDFAs) with 1-dB gain ripple across its gain bandwidth were used. A typical gain curve for such an EDFA with a total output power of +18 dB is shown in Fig. 4. This was the base-line system and the next step was to identify and assess variants of this basic configuration in terms of network viability versus cost-effectiveness. In parallel, we compared the performance of an all-optical solution against the one incorporating electronic 3R regeneration (thus forming an opaque configuration).

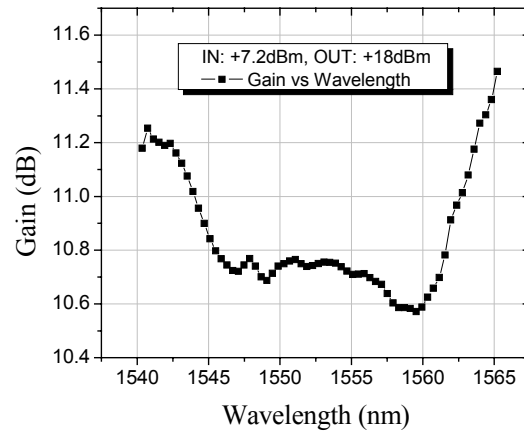


Fig. 4. Gain curve for an EDFA.

In the current work, the distance between two consecutive OPADMs was assumed to be 10 km (e.g. in Fig.1, the spans A-B, B-C, C-D etc. were assumed to be 10 km). Therefore, since the line rate for all wavelength channels was 10 Gb/s, a Dispersion Compensating Fibre (DCF) section was inserted prior to each OPADM to negate the dispersion of the previous Single Mode Fibre (SMF) section. The EDFA was a two-stage module demonstrating a 1-dB gain ripple across its gain bandwidth and it was used to compensate the losses due to the two fiber sections as well as those of the OPADM. An important aspect of the current studies is that flat passband Arrayed Waveguide Gratings (AWGs) were used in the Active configuration since it was shown in [10] that the spectral narrowing is the main limiting factor in cascading more OPADMs. The channel spacing for the Active architecture is 100 GHz while the upstream/downstream channels of the Passive case are formed when two wavebands with 100 GHz spacing are interleaved forming a 50 GHz grid in the C band. A list with the remaining most important component parameters used in these physical layer simulations is shown in Table I.

TABLE I

LIST OF THE MOST IMPORTANT SIMULATION PARAMETERS

Fibres:	SMF	DCF
Span between consecutive nodes	10 km	2 km
Attenuation	0.23 dB/km	0.50 db/km
Dispersion	17 ps/(nm.km)	-85 ps/(nm.km)
Dispersion slope	0.085 ps/(nm ² .km)	-0.3 ps/(nm ² .km)
Effective area	65 μm^3	22 μm^3
Transmitters		
PRBS	2 ⁷ -1	
Ext. ratio	14 db	
Linewidth	5 MHz	

Rise time	25 ps
Mux/Demux	
Filter response	2 nd order Gaussian
3 dB bandwidth	60 GHz
Coupling loss	6 dB
Crosstalk	-33 dB
Band Mux/Demux	
Filter response	trapezoidal
Bandwidth	400 GHz
Coupling loss	-3.5 dB
SOAs	
Small signal gain	14 dB
Input sat. power	-6 dBm
Noise figure	11 dB
EDFAs (Ideal)	
Tot. output power	18 or 23 dBm
Noise figure	5.5 dB

In the DAVID network architecture, the OPR could be connected to the MAN rings in either an optically transparent or an opaque mode. The former is made feasible thanks to 2R regenerative capabilities provided by Wavelength Converters (WC) at the OPR's output. This all-optical WC exploits the Cross Phase Modulation (XPM) of a Mach-Zender interferometer with two SOA's. The WC has been simulated using a static model in a commercial simulation tool (VPItransmissionMaker) and their regenerative properties were studied by means of the corresponding non-linear transfer function. The result of implementing a static model is that the chirping effects introduced by the converters (which interfere in either a positive or negative way on dispersion compensation) are ignored. A detailed discussion of the regenerative capability of these converters is presented in [12].

Detailed simulation studies have shown that both transparent and opaque solutions allow cascading a considerable number of OPADMs. In fact, the transparent option further improves the end-to-end performance due to the higher extinction ratio provided by the wavelength converters (22 dB versus the 14 dB when an Integrated Laser Modulator, ILM, is used). This is an important finding with respect to the overall MAN cost, since it allows removing an additional O/E/O stage. Therefore, the all-optical solution is implemented in the remaining simulations. For the study of node cascadeability, the Q-factor was used as a merit function. Specifically, a path is considered acceptable if $Q \geq 7$ (i.e. the Bit Error Rate, BER < 10^{-9}) for all WDM channels in the comb. In the commercial tool used for the simulations, the Q-factor estimations are based on the method discussed in [13].

In that method, the implicit assumption is that the induced noise has Gaussian statistical properties. This approximation is fairly accurate, since in our system the predominant degradation sources are Amplified Spontaneous Emission (ASE) noise, thermal noise and fibre non-linearities which can also be treated as Gaussian perturbations. Based on these assumptions, the Q-factor is calculated as a function of the cascaded OPADMs for the most depleted channel. This is shown in Fig. 5 for the downstream traffic for an EDFA with (a) +18 dBm and (b) +23 dBm output power. As shown in Fig. 5a, 16 OPADMs can be cascaded using both the Passive and Active configurations indicating that a ring of 160 Km in length is feasible. [If we aim at a ring capacity of 320 Gb/s, this implies an OPADM capacity of 20 Gb/s.] Also, Fig. 5a reveals the significant differences between Passive and Active structures in terms of physical layer properties. Indeed, the reason for this enhanced performance of the Active case (compared to the Passive one) is that the SOAs in the WS of the Active configuration compensate for the gain ripple when certain system conditions are met, as explained next. The SOA modules operate in the linear regime whereas the input power is at the order of -5 dBm and, thus, the data-pattern gain transients are of no concern. In addition when adding/dropping channels, the SOA gates do not induce any effective degradation on the system performance because their ON/OFF switching time is approximately 1 ns which is much less than the 50 ns guard band between successive time slots employed in the DAVID project. Therefore in the simulations, the SOA modules were implemented as static gain elements (black box) and the dynamic phenomena were ignored. Further, when a +23 dBm EDFA was used, both Passive and Active configurations can cascade more than 20 OPADMs. However, this alternative comes at a higher cost due to the higher cost of these EDFAs.

(a)

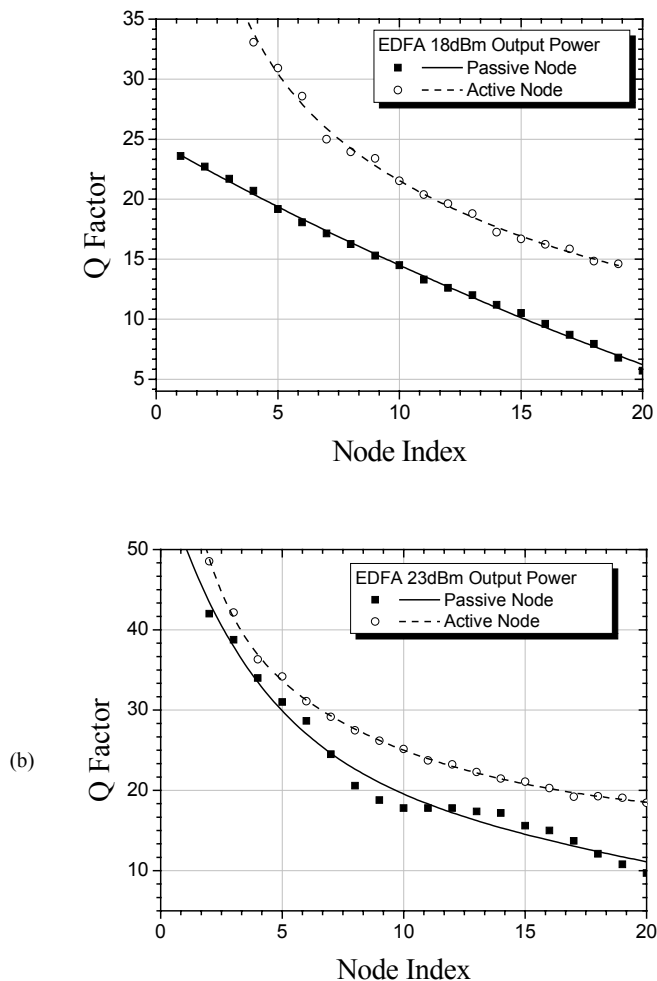


Fig. 5. Q-factor versus Node index assuming 2-stage EDFAs with 1-dB gain ripple of (a) +18 dBm and (b) +23 dBm total output power

Further, an interesting prospect is to assess the viability of configurations where one EDFA is used for compensating the losses of two OPADMs. The EDFAs under consideration have again a gain ripple of 1-dB. Detailed analysis has shown that for the Active configuration this option is not a viable alternative since the very fast Optical Signal-to-Noise Ratio (OSNR) degradation results in rapid Q-factor deterioration. Indeed, OPADM cascadeability is limited to only 7 nodes if an EDFA with +23dBm total output power and even less OPADMs are cascaded with lower power EDFAs. On the other hand, the designer has the option of placing the EDFA either at the OPADM input or to insert the node between the two EDFA stages when the Passive configuration is used. The former scheme leads to systems where the downstream channels have power levels that are very close to receiver sensitivity and this option is not considered further.

When the OPADM is inserted between the two stages of the EDFA, the dropped channels in downstream direction have sufficient power for a high OSNR. The Q-factor evolution with OPADM cascade is shown in Fig. 6 for an EDFA with (a) +18 dBm and (b) +23 dBm output power. The Q-factor is demonstrated for two channels i.e. for that located at the shorter

wavelength side of the spectrum as well as for the one with the worst performance. As Fig. 6 shows, no more than 10 OPADMs can be cascaded in both cases. When a lower power EDFA is used, the main limiting factor is the power depletion whilst in Fig. 6b the small cascadeability is caused by fibre non-linearities.

The oscillatory behavior of the shorter wavelength channel for the +23 dBm output power EDFA is a direct consequence of the fact that this channel is power-budget limited. This is not the case of the channel with the worst performance for which the Q-factor is dropping in a monotonous way.

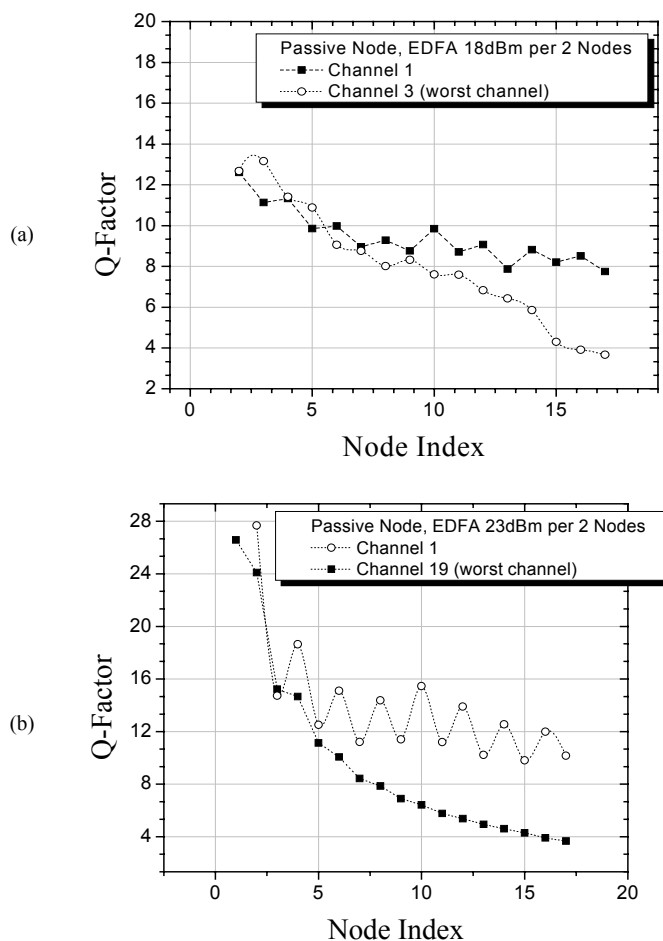


Fig. 6. The Q-factor for 1 EDFA per 2 Passive nodes for a) + 18 dBm b) +23 dBm amplifiers.

IV. COST-EFFECTIVENESS OF OPS FOR THE MAN

Now that we have assessed the physical viability of the proposed OPS architectures, we analyze whether the Active and Passive DAVID architectures are competitive with respect to more traditional approaches, namely SDH ring, star Ethernet, and RPR. Results in this section received important inputs, in terms of traffic scenarios and of network architectures, by

manufacturer and operator members of the DAVID project, which made available their internal confidential information to all partners.

A. Benchmarked solutions

The methodology consisted of fixing an initial traffic matrix and applying it to the different network architectures. Through computer simulation and/or analytical models, we determined the resources required in each network architecture (number of transceivers, number of wavelengths, number of optical amplifiers, etc.) to have similar performance (packet loss rate, delay, jitter). Finally, we used CAPEX and OPEX models in order to obtain cost benchmarking using the results of the dimensioning studies.

We restrict the study to a common network scenario with one Hub and 16 nodes distributed over a 100-km ring network. Four different node types are considered: 1 server node, 2 big nodes, 4 medium nodes, and 9 small nodes. We also considered three different mean traffic volumes: 20 Gbit/s (20G), 40 Gbit/s (40G) and 80 Gbit/s (80G). In addition we fixed the ratio between the up- and downstream traffic in the network and the number of nodes per type on the ring. This is summarized in Table II. Finally, we considered 55% of the total generated traffic coming from the backbone through the gateway, while 80% of the traffic generated at the nodes was destined to the gateway. The network characteristics considered in this study were chosen to reflect typical metro scenarios encountered by operators. The ring length of 100 km was chosen to be compatible with the node cascadeability constraints derived in the previous section III, while the number of nodes was chosen to match the limiting size of SONET/SDH rings. The diversity in the node types and their respective traffic volumes are believed to be representative for mid term metro networks. For longer term approaches, we extended the capacity per ring up to 160Gbit/s or 320Gbit/s, while also increasing the number of rings to reach a total capacity close to one Tbit/s (protected).

TABLE II
NODE TYPE AND TRAFFIC ASSUMPTION

Node type	Quantity	Upstream traffic	Downstream traffic
Server	1	20.00%	2.40%
Big	2	3.20%	8.40%
Medium	4	1.60%	4.80%
Small	9	0.80%	2.40%
Total	16	40.00%	60.00%

To compare the DAVID approaches (shown in Fig. 2 and Fig. 3) with the classical Ethernet, RPR and SDH approaches, we adopted the ADM node structures shown in Fig. 7. For the Ethernet solution (Fig. 7a), we considered a star topology where each

access node was connected directly to a central Hub through an unshared point-to-point fiber connection (doubled for protection). For both the SDH (Fig. 7b) and RPR (Fig. 7c) cases, we considered an opaque structure: optical Multiplexers (MUX) and Demultiplexers (DMUX) filter the optical channels which correspond to parallel rings terminated at each node. In the SDH approach, a single Cross-Connect (XC, switching at the STM-1 or STM-4 granularity) allows to connect to multiple rings as well as to provide add/drop access. The Hub in this case also is an SDH Cross-Connect (again switching at the STM-1 or STM-4 granularity) terminating/generating all wavelengths of the rings and of the gateway. To achieve protection capability, this structure is doubled. By nature, RPR relies on a single physical ring topology. To provide access to multiple wavelengths, multiple RPR chips are provided. Interconnection between the various RPR rings is achieved through an IP/MPLS Router, which also provides add/drop access to each of the thus stacked wavelength-rings. At the Hub, RPR interfaces are needed for all wavelengths and for connecting the gateway. The RPR architecture inherently has protection capabilities, since each physical ring is in fact composed of two counter-rotating rings [14]. As in Fig. 2 and Fig. 3, all node architectures include DABs to aggregate the data traffic coming from/going to the client layer.

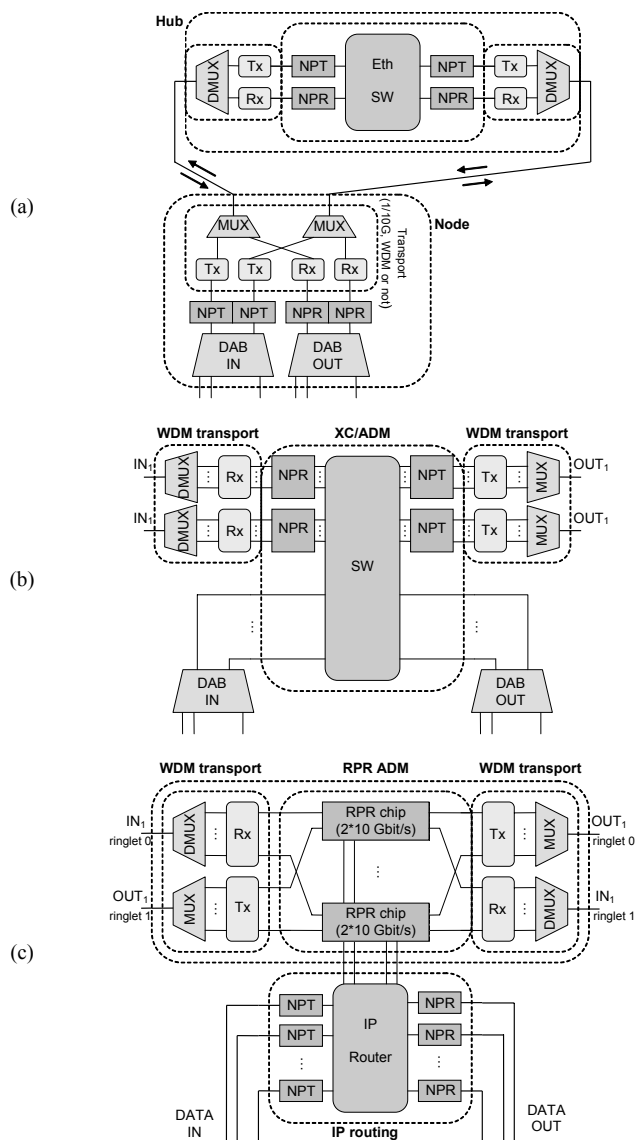


Fig. 7. OPADM node structures. (DMUX: wavelength demultiplexer; MUX: wavelength multiplexer; NPR: network processing receiver; NPT: network processing transmitter; SW: STM-1/STM-4 switch; Eth SW: Ethernet switch; XC: cross-connect; other acronyms: same meaning as in Fig. 2).

a) Point-to-point Ethernet Hub + Node. b) SDH node. c) RPR node.

B. Resource dimensioning

Taking into account the functionality and limitations of each network architecture, we performed benchmarking studies dimensioning the capacity required in each node and at the Hub to obtain similar performance. For this study we did not include any consideration of protection.

In Fig. 8, we show the node capacity (in Gbit/s) required in each metro solution considering the three traffic volumes, while Table III illustrates the needs in terms of transport resources: fibers (including the connection between the Hub and the gateway), wavelengths (either 1 Gbit/s or 10 Gbit/s channels), and transceivers (either 1 Gbit/s or 10 Gbit/s TRx).

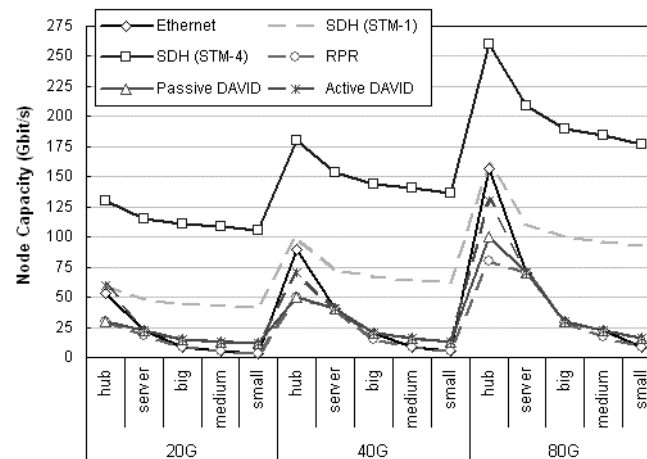


Fig. 8. Node capacity (Gbit/s) required in the different network architecture for the three traffic volumes.

Fig. 8 offers important results concerning the required resources as well as the scalability of the architecture when the traffic increases. In the SDH case, the major part of the ADM size is used for transit traffic (hence the smaller relative differences in required capacity between node types) which causes over-dimensioning. This stems from the fact that at least one circuit must be established between each source-destination pair in the network. This effect can be considerably reduced by using SDH circuits just between nodes and the Hub in a star topology rather adopting the ring-approach. In contrast, for the Active DAVID architecture, the required node capacity is nearly optimal thanks to the flexible design and the optical by-pass capability. Nonetheless, the waveband concept (which avoids the need of a full 32-wavelength selector in each ring node) imposes an over-dimensioning of the Hub. The packet-based Passive DAVID, RPR and Ethernet solutions are very similar in terms of dimensioning. Ethernet has a slight gain in nodes due to the possibility to use low bit rate interfaces, but a drawback at the Hub due to the non-shared transport resources and the star topology. For instance, it requires 44 fibers for the 40G scenario. From Fig. 8, the RPR solution seems the better one since all nodes, as well as the Hub require less capacity with respect to the other solutions. Nevertheless, the opaque structure of the RPR forces a high number of transceivers as show in Table III.

TABLE III
TRANSPORT RESOURCES REQUIRED IN THE DIFFERENT ARCHITECTURES

Equipment	Ethernet	SDH (STM-1)	SDH (STM-4)	RPR	Passive DAVID	Active DAVID
Fiber	31	2	2	3	2	2
1G Ch.	46	0	0	0	0	0
20G 10G Ch.	6	6	13	3	6	6
1G TRx	46	0	0	0	0	0
10G TRx	6	72	176	21	24	33
Fiber	44	2	2	3	2	2
1G Ch.	60	0	0	0	0	0
40G 10G Ch.	12	10	18	5	10	7
1G TRx	60	0	0	0	0	0
10G TRx	12	110	231	40	29	36
Fiber	72	2	2	3	2	2
1G Ch.	54	0	0	0	0	0
80G 10G Ch.	26	16	26	8	19	13
1G TRx	54	0	0	0	0	0
10G TRx	26	167	307	61	41	50

C. CAPEX comparison

We carried out an extensive CAPEX analysis based on the resource requirements highlighted in the dimensioning studies (physical and logical) of each architecture, based on component costs obtained by confidential means and market survey. Due to space and confidentiality limitations, we cannot provide details of our cost assumptions, which were based on input provided by manufacturers and operators in the DAVID project. We summarize the outcome of the CAPEX analysis in Table IV, where the costs are counted relative to the CAPEX for the Passive DAVID architecture. RPR is the cheapest solution only for the initial capacity: when increasing network capacity, the optical transparency provided by the Passive optical architecture enables to obtain lower CAPEX. Indeed, the Passive DAVID solution is quite competitive even for low traffic volumes, being the second cheapest solution behind RPR for the 40G traffic scenario and the cheapest solution for the 80G traffic scenario. The Ethernet and SDH solutions are in most cases not highly competitive, in both cases due to the non-sharing of resources.

TABLE IV
CAPEX COMPARISON

	Ethernet	SDH (STM-1)	SDH (STM-4)	RPR	Passive DAVID	Active DAVID
20G	-15%	+10%	+135%	-28%	0%	+58%
40G	+14%	+38%	+167%	-13%	0%	+50%
80G	+23%	+65%	+189%	+19%	0%	+45%

The Active DAVID solution pays, with the initial assumptions (limited capacity), for the complexity of the OPADM. In addition, the traffic matrix with a high proportion of extra-ring traffic (80%) is clearly a disadvantage for the Active DAVID solution which can not strongly exploit the optical space re-use mechanism.

It is important to note that the limited capacity penalizes the use of an optical Hub for both Passive and Active DAVID networks. Therefore, we benchmarked an architecture similar to the Passive DAVID solution but using an Ethernet switch at the Hub (like DBORN [15]). With respect to the Passive DAVID CAPEX, we obtained -26%, -24% and -22% for the 20G, 40G, and 80G scenario, respectively. These results clearly indicate that this solution is more appropriate for a first introduction of optical packets in metro networks.

Additional studies have been carried out considering the sensitivity of the costs. For instance, the RPR solution is quite dependent on the cost of the transceivers: dividing it by 2, the CAPEX cost is closing to the cost of passive DAVID solution; multiplying by 2, the Ethernet becomes the second cheapest solution. On the other hand, all architectures benefit from the reduction of the optics cost, but naturally the Active and Passive DAVID solutions profit most. By reducing the cost of advanced optics by a factor of 4, the Active DAVID solution becomes competitive and close to RPR in the 80G scenario.

Finally, we extended the initial traffic matrices up to 1Tbit/s scenarios (with 160G or 320G on 4 or 2 rings, where each ring should be further doubled for protection) implementing correcting factors to the initial component cost assumptions to take into account some optics cost reduction (foreseen at production of higher volumes for optical components) and a non-linear cost for electronic Tera-routers (due to higher complexity). Under these assumptions, the Passive DAVID shows the best CAPEX value (see the CAPEX comparison values in Table V). Further improvements of the active DAVID solution are made possible with the improvement of advanced optical components such as integrated fast tunable lasers. Replacing the combination of a laser array with a SOA-array and multiplexer (as presented on Fig. 3) by an integrated tunable laser having an interesting cost target, the active DAVID solution has only an extra cost of 14%, despite a traffic matrix in favor of the Passive solution.

TABLE V
CAPEX COMPARISON

	Ethernet	RPR	Passive DAVID	Active DAVID
1T	+87%	+56%	0%	+25%

D. OPEX comparison

For the OPEX comparison, we adopted a common model where annual costs have been calculated as a percentage of the equipment costs. The justification of this approach is based on the fact that OPEX is related to complexity, functionality, size, power and construction of hardware, which again is related to CAPEX. We considered this a valid approach, even though it is only a first approximation and best represents a greenfield case. In other cases an operator must also consider migration issues, which can have significant influence on OPEX.

OPEX includes various operational costs, ranging from administrative costs, over service development, to network planning costs etc. We limited our comparison to costs related to network *operations* and *maintenance* mainly because other cost factors are most likely to not significantly differ between the various architectures. Also migration costs have not been accounted for.

The *maintenance costs* have been defined as all the costs related to the resolution of physical problems in the network such as fiber cuts or equipment failure. It can be calculated as the sum of *replacement costs* and the *maintenance staff costs*. The first part encompasses the cost of failed network elements and is proportional to its failure probability, while the second includes labour costs and obviously depend on the required amount of personnel.

The operational costs include all the recurrent costs which are periodically necessary for undisturbed operation. Thus, costs for electrical power are a part of the operational costs as well as the reconfiguration costs after a failure. Operational costs have been calculated as a percentage of the equipment cost (proportion cost \times equipment cost). The proportion cost factors—which are dependent on the type of equipment/component—have been empirically obtained after assessing the overall operation costs of several real metropolitan networks exploited by the operators involved in the DAVID project and considering several activities such as reconfiguration, supervision, network element database management, energy consumption, software upgrades, etc.

OPEX results for the different network architectures are depicted in Fig. 9. The cost specific values are expressed as relative to the cost of 1 fibre.km.

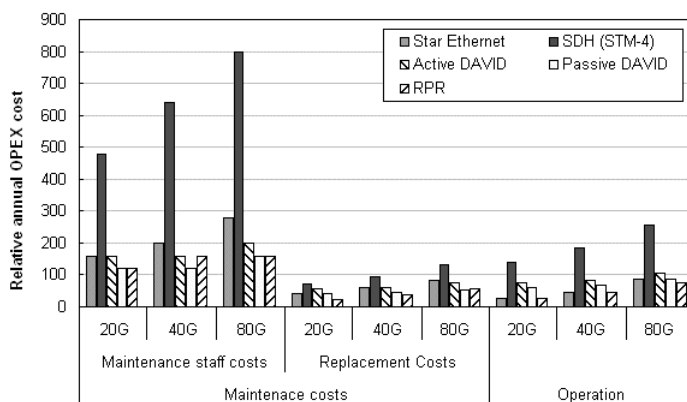


Fig. 9. Relative annual OPEX cost comparison in the different network architecture for the three traffic volumes.

The OPEX costs for the SDH ring are considerably larger than for the other scenarios, since it includes many more network elements, most of them electronics. RPR on the other hand, being the option with the fewest number of network elements, presents the lowest OPEX costs. Yet it is closely followed by the Passive DAVID, Ethernet and Active DAVID solutions.

The ratio of annual OPEX over CAPEX, shown in Fig. 10, provides an indication of OPEX vs. CAPEX over the life-cycle of the network platform. These results indicate that the solutions based on optical packet switching in general have a lower ratio than the other solutions based on electronic switching. Thus, if OPADM CAPEX is comparable to competing solutions based on traditional technology one can expect that there is a cost reduction potential in OPADM OPEX over the traditional solutions.

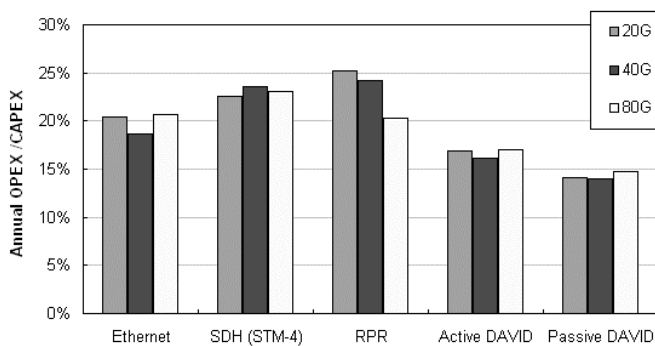


Fig. 10. Annual OPEX over CAPEX.

The relative levels of the OPEX components are shown in Fig. 11. Maintenance staff costs are the dominant factor for all solutions. However, the numbers indicate that in general the OPADM solutions lead to lower maintenance staff cost and higher replacement and operational costs considered relative to each other.

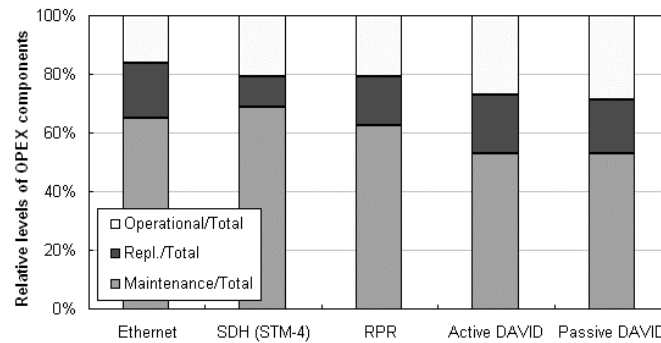


Fig. 11. Relative levels of OPEX components, averaged over the 20G, 40G, and 80G cases.

OPEX sensitivity calculations were carried out to get a better understanding on how OPEX of the different architectures depend on investment costs of various component categories as well as on Mean Time Between Failure (MTBF) of the different component categories. The aim was also to get a better understanding of OPEX cost model itself. The focus was on replacement and operations costs, assuming that maintenance staff to a large degree is proportional to replacement costs.

We observed that all solutions and in particular SDH and RPR solutions, are relatively more sensitive to changes of TRx and electronic costs than OPADM solutions are sensitive to changes in costs of advanced optics. The cost for the star Ethernet solution is characterized by higher fibre maintenance and replacement costs, and therefore was found to be relatively less sensitive to changes in TRx and electronic costs.

The operations proportion cost factors for advanced optics components—for which very little experience is available—as well as the replacement cost factors ($\sim 1/\text{MTBF}$) can easily be doubled without affecting the OPEX position of the DAVID architectures. Indeed, OPEX is highly dominated by maintenance costs (Fig. 11), and the gap with traditional technologies is quite large (Fig. 9). Thus, based on our sensitivity studies, we maintain the conclusion that the DAVID architectures achieve lower relative annual OPEX compared to the traditional metro solutions.

E. Conclusions on cost-effectiveness

From the extensive benchmarking results—whose results have been summarized above—, and despite all uncertainties of market analyses and forecasts, we can foresee a possible introduction scenario of the different metro technologies with respect to the required capacity and the traffic repartition. Fig. 12 depicts this scenario, whose tendencies could be summarized as follows:

- With low capacity (few 10's Gbit/s), two advantageous solutions can be identified: the star Ethernet (possibly with WDM to share fibre resources) when the ratio of intra-ring traffic is low, whereas RPR appears the most optimized solution thanks to space re-use capability.

- At a short/medium term with increasing access bit rate and resulting metro capacity in the range of 10's to few 100's of Gbit/s, passive optical ring structure with an electrical Hub is well suited, as in the DBORN architecture proposed by Alcatel [15]. Due to the lack of transparency, RPR requires a high amount of transceivers and filtering ports on the ring which makes the solution less competitive.
- At a longer term, under the assumption of a strong introduction of high bit rate access networks (FTTx, GPONs), the capacity in the metro can reach 100's of Gbit/s to 1 Tbit/s. In this case, the two DAVID solutions become competitive, thanks to the optical transparency both at the OPADM and Hub levels.

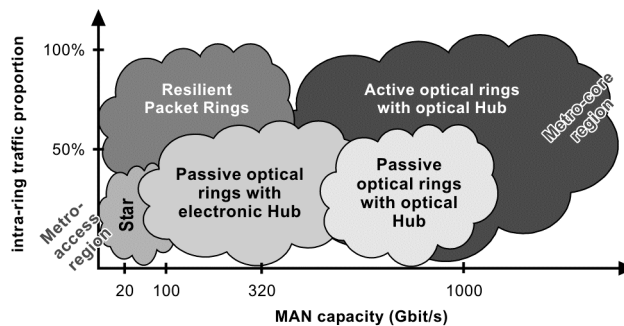


Fig. 12. Possible introduction scenario of the different metro technologies.

V. PERFORMANCE EVALUATION OF THE DAVID MAN ARCHITECTURES

A. Assessment of the space reuse and waveband concepts

In the previous section, we have compared the OPS architectures proposed within DAVID against competing metro alternatives. We will now focus on the Passive and Active OPS architectures only, but consider a broader range of traffic matrices [16]. In particular, we will evaluate the space reuse concept and the waveband approach of the Active architecture as sketched in Fig. 3. The approach taken was to use a heuristic planning algorithm to design a MAN ring network able to carry a given amount of traffic between a given set of nodes V . The demand is given as a matrix D , where $D(i,j)$ denotes the bandwidth required between OPADMs i and j .

The cost indicators used here are the following: (i) Tx/Rx capacity: the total number of Tx/Rx elements used, summed over all OPADMs, (ii) link capacity: the number of wavelengths effectively used per link, summed over all physical links, (iii) nr. of lambdas: the number of wavelengths used per ring, summed over all rings. The first criterion is an indicator of the OPADM costs, while the last will impact the Hub dimension and thus its cost. To assess the resource requirements of the OPADM architectures, we covered the following demand patterns: (i) Uni: a uniform demand pattern, where between each two OPADMs

a bandwidth d needs to be set-up ($D(i,j)=d$); (ii) Serv: there is one server node s , which dominates the demand matrix ($D(i,s)=D(s,i)=2d$, other $D(i,j)=d$); (iii) Neigh3: each node only communicates to 3 other nodes ($D(i,i+1)=D(i,i+2)=D(i,i+3)=d$, rest is zero); (v) David: the demand matrix used in the aforementioned benchmarking studies.

The main difference between the Active and Passive architectures from a conceptual point of view is the space reuse potential of the Active structure. Fig. 13 presents dimensioning results for Passive and Active architectures with wavebands of a single wavelength. (Note that $B=1$ amounts to having no waveband concept; $B>1$ is discussed in the next paragraph.) From a Tx/Rx cost perspective, we conclude that the Active approach needs more Tx/Rx capacity. The reason is that to allow space reuse, the receiver and transmitter have to be able to access the same wavelength, which sometimes requires an extra Tx/Rx (cf. $B=1$ means no tunability in the architecture as given in Fig. 3). The space reuse concept proves useful when the CAPEX of the MAN is dominated by the link capacity, or the number of wavelengths per ring. This is due to the fact that there is no spectral separation for up- and downstream, and the space reuse capability allows for better sharing of the available bandwidth among different demands. Note that, in several scenarios, using extra optical bandwidth on the fibers, without requiring extra switching nor extra electronic capacity, leads to marginal cost increases.

A second aspect in which the Active and Passive structures differ is the waveband concept. In the space reuse assessment, we used wavebands of a single wavelength (ie. no tunability in the OPADMs). In this section we study the impact of introducing the waveband concept, again from a network dimensioning point of view. We compare the Active nodes with $B=1$ versus $B=4$ in Fig. 14. The advantage of the band concept is that Tx/Rx capacities can be somewhat reduced. Yet, when CAPEX is dominated by link capacities, the band concept is not useful, since it heavily increases the number of wavelengths used, indicating that spatial reuse opportunities within bands are limited. This stems from the fact that the architecture is assumed to allow only a single Tx/Rx per band per OPADM. (Note that in Section III, we have assessed also the viability of all OPADMs to add/drop any channel.)

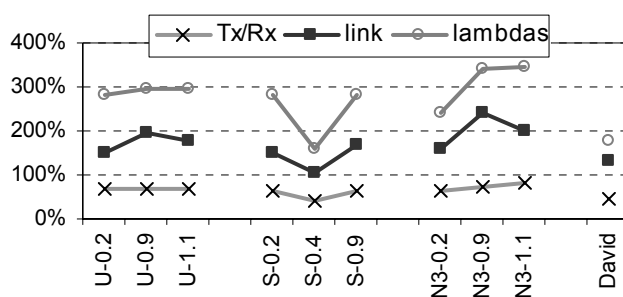


Fig. 13. Cost ratio Passive/Active; x-axis labels denote demand (U=Uni, S=Serv, N3=Neigh3; the number after the dash is the value of d).

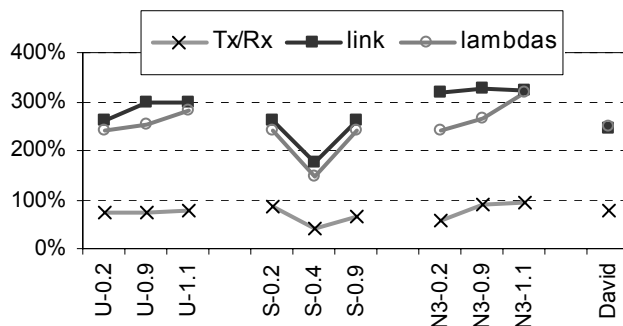


Fig. 14. Cost ratio (bands, B=4)/(no bands, B=1); x-axis labels denote demand.

B. MAC operation and packet scheduling

The above comparisons between different architectures are based upon a particular, static, traffic matrix: OPADM and Hub dimensioning is optimized for the given traffic matrix, and the required hardware is evaluated. However, the traffic matrix often does not behave statically; thus, in this section we fix OPADMs and Hub configurations, and then check which traffic patterns and loads can be supported by the given network architectures. We first comment on the basic DAVID MAN design choices; subsequently, performance results are introduced by a motivated description of the adopted scheduling algorithm (needed to allocate network resources).

As outlined in Section II, no optical buffering is used in the DAVID MAN architectures, neither in the OPADMs, nor in the Hub. Thus, the Hub operates as a space/wavelength switch: in every slot it provides input/output permutations, which can be either (i) wavelength-to-wavelength or (ii) ring-to-ring. In the former case (i), packets received from an input wavelength channel are forwarded to an output wavelength, connecting input/output channels in disjoint pairs through an input/output permutation. Thus contention—and therefore the need for storing contending packets—is avoided. In the latter case (ii), all packets received from an input ring are forwarded to the same output ring; again input/output rings are selected in disjoint pairs, and the number of wavelengths in each ring must be the same to avoid contentions. The input/output permutation at the Hub can be changed at every slot. This requires fast switchable or tunable components in the Hub.

The scheduling algorithm must control resource allocation (time slots and wavelengths). To reduce the complexity, a distributed approach to resource allocation is preferred. Distributed access decisions are based upon processing the control channel (through which nodes are capable of knowing if current data slots are busy or free). This channel inspection capability prevents collisions (more than one packet transmitted in the same slot of the same wavelength at the same time), but can also prevent contentions (more than one packet to be received by the same OPADM at the same time). A distributed contention

avoiding MAC protocol is possible only if the Hub operates through ring-to-ring permutations, so that all slots that will be seen in a future time slot by a receiver can be simultaneously inspected by the transmitting OPADM. Note that the possibility of distributed prevention of collisions and contentions largely improves the scalability of the network, and is an advantage of rings with respect to star topologies.

Two levels of scheduling arise in the network: the first one is performed at the Hub to allocate ring-to-ring bandwidth, whereas the second one is performed in a distributed way at each OPADM to fit packets into ring-to-ring bandwidth pipes. The problem of finding an optimal sequence of permutations at the Hub can be formalized as an optimization problem. The scheduling for a given traffic matrix can be computed at the Hub by using standard techniques based on iterated applications of approximated maximum size or maximum weight matching algorithms. An optimal solution and a number of heuristic solutions to this problem were presented and studied in [17]. The information upon which the Hub scheduling is computed is a ring-to-ring traffic matrix, which can be either estimated at the Hub by means of measurements [18], or built with explicit reservations issued by OPADMs to the Hub. Here we simply assume that the traffic matrix is known and hub scheduling is matched to the traffic matrix. It is worth mentioning that the complexity of the scheduling algorithm at the Hub is kept low since it scales with the number of rings in the network (instead of scaling with the number of network nodes).

As described in Section II, a separate control wavelength is dedicated to signaling purposes. The ring-to-ring permutation sequence generated by the scheduling algorithm at the Hub is notified to OPADMs by writing on the signaling channel the destination-ring of each slot. This signaling permits distributed access decisions. Nodes must know the busy/free state of all data slots to avoid collisions (i.e., to refrain from transmission on an already used slot), as well as the destination of packets transmitted by upstream nodes to avoid receiver contentions.

Even if the Active node architecture enables space reuse—as discussed above—to simplify the scheduling problem and to provide a fair comparison among architectures we assume in the sequel that space reuse is not exploited; thus, each slot can be used at most once. This allows us to concentrate on transceivers' tunability, and elaborate on the effect of tunability on network performance. For similar reasons, we assume that wavelength separation of up- and downstream traffic is adopted in both architectures. Thus, the only remaining difference is that the Active architecture has tunable transceivers, while the Passive has not. (Recall that a transceiver, even if it's tunable, is capable of transmitting and receiving only on one channel at a time, i.e. in a particular timeslot.)

We consider a network with 4 rings, where each ring has 16 nodes and 4+4 data wavelengths (wavelength separation of transmission and reception). Each ring is assumed to have the same length, i.e. contain the same number of slots. We examine only two traffic patterns due to space constraints. The first is a simple uniform model. The second one, named realistic, stems

from real traffic estimates and is a straightforward extension to the multi-ring scenario used for benchmarking studies. All rings are alike in terms of nodes and traffic distribution, so that we can express the probability of transmitting a packet from node i of ring k to node j of ring l as $N(i,j) \times R(k,l)$, where $N(i,j)$ is the probability of transmitting a packet from node i to node j inside any ring, and $R(k,l)$ is the probability of transmitting a packet from ring k to ring l . The node-to-node rates are suitably scaled to obtain the desired total network load.

Under uniform traffic, all nodes generate packets at the same rate and with the same destination probability. So the probability of transmitting between any two nodes is always equal to $1/64$. Under the realistic pattern, the load on all rings is the same, and at each ring four types of nodes, named server, big, medium and small, can be identified, each one with its own particular characteristics (load and location in the ring), as described in Table VI. One server node, two large nodes, four medium nodes and nine small nodes are present on each ring. Matrix R is reported below, while matrix N is equal to the matrix described in Section IV.A. About 20% of the generated traffic remains on the source ring, while the remaining 80% is distributed to outside rings according to the corresponding node transmission probabilities.

TABLE VI
NODE CHARACTERISTICS UNDER REALISTIC TRAFFIC PATTERN FOR ANY RING

Node Type	Percentage of total traffic generated by a single node	Node number
Server	50 %	4
Large	8 %	1, 7
Medium	4 %	3, 8, 11, 13
Small	2 %	2, 5, 6, 9, 10, 12, 14, 15, 16

$$\mathbf{R}_{\text{realistic}} = \begin{pmatrix} 2/10 & 8/30 & 8/30 & 8/30 \\ 8/30 & 2/10 & 8/30 & 8/30 \\ 8/30 & 8/30 & 2/10 & 8/30 \\ 8/30 & 8/30 & 8/30 & 2/10 \end{pmatrix}$$

Under the uniform traffic pattern both architectures allow almost 100% throughput to be obtained and no significant differences can be observed.

Fig. 15 plots the normalized ring throughput versus the normalized ring offered load under realistic traffic. Receivers are uniformly distributed among wavelengths in the Passive architecture. This choice is not optimal for the considered traffic pattern, but would be a reasonable choice to cope with unknown traffic patterns. When increasing the offered load, some channels

become overloaded earlier than others due to an unbalanced traffic matrix. Thus, throughput curves grow linearly with decreasing slopes until all channels become overloaded and throughput saturates (not shown in the figure).

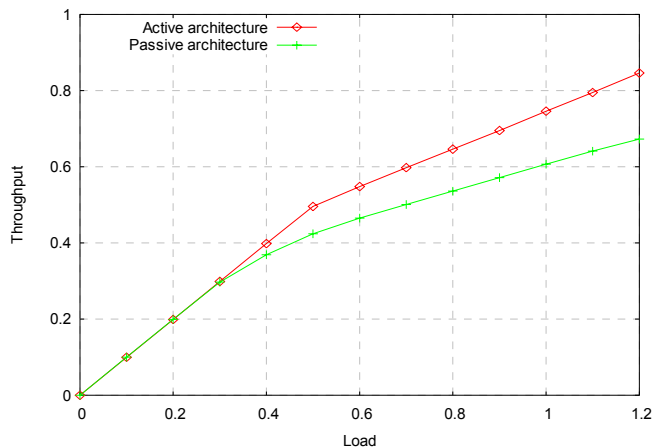


Fig. 15. Overall network throughput under the realistic traffic pattern

As expected, the Passive architecture results in overloaded channels and worse performance, due to the non-optimal allocation of transceivers to wavelength channels. Since in the Active architecture fast tunable transmitters and receivers are available, it is obvious that the former can easily adapt to variations of the traffic matrix by switching traffic from overloaded wavelengths to unloaded ones, while the latter is limited to transmitting and receiving on fixed channels. However, if the fixed transmitters and receivers of the passive architecture nodes are optimally, but still statically, allocated to wavelengths, non reported simulation results show that the observed performance degradation are almost completely recovered. One way to achieve a good allocation of transceivers, without requiring transceivers to fast-tune to wavelengths in a slot-by-slot fashion, thereby keeping low hardware requirements, can be to introduce a slow tuning capability, which permits to slowly re-allocate transceivers to wavelengths, and to adapt the network configuration to slow variations of the traffic matrix. This approach builds upon the observation that variations in the traffic matrices happen on a much longer time scale than packet-by-packet allocation, so that very fast, i.e. slot by slot, transceiver tunability may not be necessary. It is an interesting issue to critically re-think about what degree of fast tunability/switching is really required in optical packet switched networks in general. Typically QoS requirements are dictated by users' needs, and users' needs do not scale with transmission speed. This means that it may not be really necessary to keep increasing the switching speed of packet networks with the increase of line rates, but resource allocation decisions may be taken at a slower time scale, compatible with the QoS targets. It is envisaged that these considerations will be mostly important in the design of next-generation optical and electronic packet-switched networks.

VI. CONCLUSIONS

Today's and tomorrow's metro networks are characterized by dynamic traffic scenarios, both on a temporal and spatial scale. Optical packet switching exploits efficient optical switching of high capacity traffic streams with the bandwidth efficiency of a packet switching paradigm. In the European research project DAVID, two OPS ring architectures have been proposed for MAN environments. In this paper we have argued their viability from a physical performance point of view, showing that at least 16 OPADMs can be cascaded in a single ring.

Furthermore, dimensioning studies compared the proposed OPS architectures with more conservative architectures such as SDH/SONET, Ethernet and RPR. Both CAPEX and OPEX studies revealed that the OPS architectures, and especially the Passive one, are strong potential competitors. The Passive architecture was found to lead to lowest CAPEX for high traffic volumes. Accounting for reduced costs of optics as foreseen in future, the Active architecture reached lowest CAPEX. Both DAVID architectures lead to a reduction of the relative proportion of OPEX in the overall cost.

Finally, we compared the assets of Active and Passive structures from a logical performance point of view. Extensions of the dimensioning studies to a broad set of traffic matrices showed that the Active structure is particularly helpful to reduce the amount of wavelengths needed. Simulation studies with dynamically varying traffic illustrated the efficiency of a Hub scheduling algorithm and MAC. We also showed that exploiting wavelength conversion greatly improves throughput for medium to highly loaded rings.

In conclusion, we have demonstrated that a ring-based OPS architecture is a competitive and highly viable approach for future metro networks.

ACKNOWLEDGMENT

The authors would like to thank their colleagues of the DAVID project for the many fruitful discussions.

REFERENCES

- [1] P. Zhou, O. Yang, "How practical is optical packet switching in core networks?", *Proc. IEEE Global Telecommun. Conf. (GLOBECOM 2003)*, San Francisco, CA, vol. 5, 1-5 Dec. 2003, pp. 2709–2713.
- [2] C. Develder, J. Cheyns, E. Van Breusegem, E. Baert, D. Colle, M. Pickavet, P. Demeester, "Architectures for optical packet and burst switches" (Invited), *Proc. 29th European Conf. on Optical Commun. / 14th Int. Conf. on Integrated Optics and Optical Fibre Commun. (ECOC-IOOC)*, Rimini, Italy, vol. 1, 21-25 Sep. 2003, pp. 100–103.

- [3] K. Shrikhande, et al., "HORNET: A packet-over-WDM multiple access metropolitan area ring network", *IEEE J. Selected Areas in Commun.*, vol 18, no. 10, Oct. 2000, pp. 2004–2016.
- [4] A. Takada, J.H. Park, "Architecture of ultrafast optical packet switching ring network", *IEEE J. Lightwave Techn.*, vol. 20, no. 12, Dec. 2002, pp. 2306–2315.
- [5] D. Dey, A. van Bochove, A. Koonen, D. Geuzebroek, M. Salvador, "FLAMINGO: a packet-switched IP-over-WDM all-optical MAN", *Proc. European Conf. on Optical Commun. (ECOC 2001)*, Amsterdam, The Netherlands, 30 Sep. – 4 Oct. 2001, pp. 480–481.
- [6] A. Carena, V. Ferrero, R. Gaudino, V. De Feo, F. Neri, P. Poggiolini, "RINGO: A demonstrator of WDM optical packet network on a ring topology", *Proc. 6th IFIP Working Conf. on Optical Netw. Design and Modeling (ONDM 2002)*, Torino, Italy, 4-6 Feb. 2002, pp. 183–197.
- [7] L. Dittman, et al., "The European IST project DAVID: a viable approach towards optical packet switching", *IEEE J. Selected Areas in Commun.*, vol. 21, no. 7, Sept. 2003, pp. 1026–1040.
- [8] S. Yao, S. J. Ben Yoo, B. Mukherjee S. Dixit, "All-optical packet switching for metropolitan area networks: Opportunities and challenges", *IEEE Commun. Mag.*, vol. 39, no. 3, March 2001, pp. 142–148.
- [9] F. Callegati, C. Devellder, W. Cerroni, M. Pickavet, G. Corazza, P. Demeester, "Scheduling algorithms for a slotted packet switch with either fixed or variable length packets", *Photonic Netw. Commun. (PNET)*, vol. 8, no. 2, Sep. 2004.
- [10] A. Stavdas, S. Sygletos, M. O'Mahoney, H.L. Lee, C. Matrakidis, A. Dupas, "IST-DAVID: concept presentation and physical layer modeling of the metropolitan area network", *IEEE J. Lightwave Techn.*, vol. 21, no. 2, Feb. 2003, pp.372–383.
- [11] B.B. Mortensen, M.S. Berger, "Optical Packet Switched Demonstrator", *Proc. Int. Conf. on Optical Internet & Photonics in Switching (COIN-PS 2003)*, Jeju Island, South Korea, Jul. 2002.
- [12] J.-Y. Emery, M. Picq, F. Poingt, F. Gaborit, R. Brenot, M. Renaud, B. Lavigne, A. Dupas, "Optimised 2R all-optical regenerator with low polarisation sensitivity penalty (<1 dB) for optical networking applications", *Proc. Optical Fiber Commun. Conf. and Exhibit (OFC 2001)*, Anaheim, CA, vol. 1, pp. MB4-1 – MB4-3.
- [13] N.S. Bergano, F.W. Kerfoot, C.R. Davidson, "Margin Measurements in Optical Amplifier Systems", *IEEE Photonics Techn. Lett.*, vol. 5, no. 3, May 1993, pp. 304–306.
- [14] F. Davik, M. Yilmaz, S. Gjessing, N. Uzun, "IEEE 802.17 resilient packet ring tutorial", *IEEE Commun. Mag.*, vol. 42, no. 3, March, 2004, pp. 112–118.
- [15] N. Le Sauze, et al., "A novel, low cost optical packet metropolitan ring architecture", *Proc. European Conf. on Optical Commun. (ECOC 2001)*, Amsterdam, The Netherlands, 30 Sep. – 4 Oct. 2001, vol.6, pp. 66–67.
- [16] R. Van Caenegem, C. Devellder, E. Baert, D. Colle, M.Pickavet, P. Demeester, "Architectures for OPS metro rings: comparing active versus passive nodes / A dimensioning point of view", *Proc. ONDM 2004 (Optical Network Design and Modelling)*, IFIP TC-6 Working Conference, Gent, Belgium, 2-4 Feb. 2004, pp. 101–120.
- [17] A. Bianco, J. Finochietto, E. Leonardi, F. Marigliano, P. Mitton, F. Neri, L. Quarello, "Multiclass resource allocation in interconnected WDM Rings", *Proc. ONDM 2003 (Optical Network Design and Modelling)*, IFIP TC-6 Working Conference, Budapest, Hungary, 3-5 February 2003
- [18] A. Bianco, G. Galante, E. Leonardi, F. Neri, "Measurement Based Resource Allocation for Interconnected WDM Rings", *Photonic Netw. Commun. (PNET)*, vol. 5, no. 1, Jan. 2003, pp. 5–22.

C. Devellder received the M.Sc. degree in computer science engineering and a Ph.D. in electrical engineering from Ghent University (Gent, Belgium), in July 1999 and December 2003 respectively. From October 1999 on, he has been working in the Department of Information Technology (INTEC), at the same

university, as a Researcher for the Fund for Scientific Research-Flanders (FWO-V), in the field of network design and planning, mainly focusing on optical packet switched networks. As such, he has been involved in e.g. the European IST-projects DAVID and STOLAS, as well as a national research project on "Optical Networking and Node Architectures". In January 2004, he joined OPNET Technologies.

A. Stavdas received the B.Sc. degree in physics from the University of Athens, Athens, Greece, and the Ph.D. degree from the University College of London, London, U.K., in the field of wavelength routed WDM networks. Currently, he is Principal Researcher of ICCS, NTUA, and he is heading the Optical Networking Group of NTUA (<http://www.ong.ece.ntua.gr>). He is the author or co-author of over 50 journal publications and conference articles. He worked in a number of European Projects like IST-DAVID, IST-OLYMPIC, AC069 COBNET, and AC050 PLANET. He also led projects funded by industry (Nortel Networks and Thermophotonics). Current interests include physical layer modeling of optical networks, optical packet switching, ultra-high capacity end-to-end optical networks, OXC architectures, and WDM access networks.

Dr. Stavdas served as Chairman of the Optical Network Design and Modeling Conference (ONDM 2000) and as a Member of the Technical Committee for Conferences like GLOBECOM and Photonics in Switching.

A. Bianco is Associate Professor at the Electronics Department of Politecnico di Torino, in Italy. He was born in Torino, Italy, in 1962. He holds a Dr.Ing. degree in Electronics Engineering since 1986 and a Ph.D. in Telecommunications Engineering since 1993, both from Politecnico di Torino. He has co-authored over 100 papers published in international journals and presented in leading international conferences in the area of telecommunication networks. His current research interests are in the fields of protocols for all-optical networks and switch architectures for high-speed networks. He participated to the technical program committees of several conferences, including IEEE Infocom 2000, QoS-IP (Quality of Service in Multiservice IP Networks) 2001, IFIP ONDM (Optical Network Design and Modelling) 2002, 2003, and 2004, and Networking 2002 and 2004. He was technical program co-chair of the HPSR (High Performance Switching and Routing) 2003 workshop. He was involved in several European (ACTS-SONATA, IST-DAVID), and Italian projects (EURO, IPPO, WONDER) on optical networks and switch architectures.

D. Careglio received the M.Sc. in Electrical Engineering from both the Universitat Politècnica de Catalunya (UPC), Spain, in 2000 and the Politecnico di Torino, Italy, in 2001. He is Assistant Professor at the Department of Computer Architecture of UPC and he is currently working toward the Ph.D. degree in the same university. He is a member of the Advanced Broadband Communications Centre of UPC (<http://www.ccaba.upc.es>). He has recently been involved in ACTS-SONATA, IST-LION and IST-DAVID European projects. His research interests are in the fields of all-optical networks with emphasis on MAC protocols, QoS provisioning, and traffic engineering.

H. Lønsethagen received a B.sc in 1987 in electrical engineering and computer science from University of Colorado, Boulder, USA. He was granted a M.sc. from the Norwegian Institute of Technology in 1988. Since 1990, with Telenor R&D, he has been working with telecom management in particular network level management of transport networks. From 1994-2000 he has been involved with activities related to systems integration and distributed systems frameworks, architectures and middleware such as RM-ODP and TINA. He has participated in Eurescom projects (P109, P924) and IST DAVID. Since 2000 his research activities have been related to traffic engineering and network control and management of packet-based networks and intelligent optical networks also including techno-economic analysis.

J.P. Fernandez-Palacios Giménez was born in Hellín (Spain) in February of 1975. In June 2000 he graduated with a degree of Telecommunications Engineer from Polytechnic University of Valencia where he carried out his final project working on the simulation of wavelength converters. In September 2000 he joined Telefónica I+D where he has been working on the analysis and evaluation of new technologies either in access or backbone networks, likewise he has participated in European projects such as Eurescom P1014 TWIN (Testing WDM IP Networks), IST DAVID (Data and Voice Integration over DWDM) and other Telefonica internal projects related to the development of optical networks in Spain and Latin America.

R. Van Caenegem received the M.Sc. degree in electrical engineering, specialized in micro- and opto-electronics from Ghent University, Gent, Belgium, in 2003. From August 2003 on, she has been working in the Department of Information Technology, at the same university.

S. Sygletos received the diploma of Electrical Engineering and Computer Science from the Faculty of Electrical and Computer Engineering of the National Technical University of Athens in 2000, with specialization in Telecommunications. He is currently working towards his Ph.D degree in the field of optical telecommunications. His current research interests include physical layer modelling of optical networks, all optical signal regeneration and fibre transmission issues

He has been involved in IST-DAVID and IST-OLYMPIC European projects as well as in a number of national research projects.

F. Neri is Full Professor in the Electronics Department, Politecnico di Torino, Turin, Italy. He received his Dr. Ing. and Ph.D. degrees in Electrical Engineering from Politecnico di Torino in 1981 and 1987, respectively. His research interests are in the fields of performance evaluation of communication networks, high-speed and all-optical networks, packet switching architectures, discrete event simulation, and queuing theory. He has co-authored over 150 papers published in international journals and presented in leading international conferences.

Fabio Neri leads a research group on optical networks at Politecnico di Torino. He has recently been involved in several European and Italian projects on WDM networks. He is the coordinator of the European 6FP Network of Excellence "e-Photon/One" on optical networks, which involves 38 European research institutions, operators and manufacturers.

Fabio Neri served several IEEE journals and conferences, including IEEE Infocom and IEEE Globecom. He was general co-chair of the 2001 IEEE Local and Metropolitan Area Networks (IEEE LANMAN) Workshop, and of the 2002 IFIP Working Conference on Optical Network Design and Modelling (ONDM). He serves in the editorial board of IEEE/ACM Transactions on Networking.

J. Solé-Pareta was awarded his Master's degree in Telecommunication Engineering in 1984, and his Ph.D. in Computer Science in 1991, both from the Universitat Politècnica de Catalunya (UPC).

In 1984 he joined the Computer Architecture Department of UPC. Since 1992 he has been an Associate Professor with this department. He is co-founder and member of the Advanced Broadband Communications Centre of UPC (<http://www.ccaba.upc.es>). His current research interests are in Broadband Internet and High Speed and Optical Networks with emphasis on traffic engineering, traffic characterisation, traffic management, MAC protocols and QoS provisioning. He has participated in INFOWIN, MICC and IMMP ACTS projects, and in LION, DAVID and LONG IST projects. Within the VI Framework Program he is participating in NOBEL (IP-project) and in e-Photon/One (Network of Excellence).

M. Pickavet received an M.Sc. and Ph.D. degree in electrical engineering, specialized in telecommunications, from Ghent University in 1996 and 1999, respectively). Since 2000, he is professor at Ghent University where he is teaching telecommunication networks and algorithm design.

His current research interests are related to broadband communication networks (WDM, IP, (G-)MPLS, OPS, OBS) and include design, long-term planning and routing of core and access networks. Special attention goes to Operations Research techniques that can be applied for routing and network design. In this context, he is currently involved a.o. in the European IST projects "Data and Voice Integration over DWDM" (DAVID), "Switching Technologies for Optically Labeled Signals" (STOLAS), "All-optical Label Swapping employing Optical Logic Gates" (LASAGNE) and the Network of Excellence on optical networks ePHOTON/ONe. He has published several international publications, both in journals (e.g. IEEE JSAC, IEEE Comm. Mag., Eur. Trans. on Telecommunications, Photonic Network Communications, ...) and in proceedings of conferences.

N. Le Sauze obtained his M. Sc. in Physics from the Université de Bretagne Occidentale, Brest, France in 1996, and his engineer diploma of telecommunication from the Ecole Nationale Supérieure des Télécommunications de Bretagne, Brest, France in 1998. He joined Alcatel CIT in 1998 to work on research activities in optical packet networking, and since 2001, he is leading a study on optical packet metropolitan networks. Through these works, he developed a strong background on optical packet networks, including optical architecture design, logical performance evaluation and benchmarking studies. He previously contributed to the French RNRT ROM and European IST DAVID projects.

P. Demeester is professor at the Ghent University. His current interests are related to broadband communication networks (IP, G-MPLS, optical packet and burst switching, access and residential, active, mobile, CDN, grid) and include network planning, network and service management, telecom software, internetworking, network protocols for QoS support, etc. He published over 250 journal or conference papers in this field.