Benchmarking and viability assessment of optical packet switching for metro networks

Benchmarking and Viability Assessment of Optical Packet Switching for Metro Networks

C. Develder, A. Stavdas, *Member, IEEE*, A. Bianco, D. Careglio, H. Lønsethagen, J. P. Fernández-Palacios Giménez, R. Van Caenegem, S. Sygletos, Fabio Neri, J. Solé-Pareta, M. Pickavet, N. Le Sauze, and P. Demeester, *Senior Member, IEEE*

***Abstract*—Optical packet switching (OPS) has been proposed as a strong candidate for future metro networks. This paper assesses the viability of an OPS-based ring architecture as proposed within the research project DAVID (Data And Voice Integration on DWDM), funded by the European Commission through the Information Society Technologies (IST) framework. Its feasibility is discussed from a physical-layer point of view, and its limitations in size are explored. Through dimensioning studies, we show that the proposed OPS architecture is competitive with respect to alternative metropolitan area network (MAN) approaches, including synchronous digital hierarchy, resilient packet rings (RPR), and star-based Ethernet. Finally, the proposed OPS architectures are discussed from a logical performance point of view, and a high-quality scheduling algorithm to control the packet-switching operations in the rings is explained.**

***Index Terms*—Optical packet switching, wavelength-division multiplexing, metropolitan area networks (MANs), performance, medium access control (MAC).**

I. Introduction

DESPITE the recent economic malaise, the demand for telecommunication services continues to grow steadily. Even though this growth may have been overenthusiastically 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 fiber technology. Indeed, the use of wavelength-division multiplexing (WDM) offers massive bandwidth through the parallel transmission of high-bit-rate channels onto the same fiber 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 (ASONs), enabling the automated setup and tear-down of so-called lightpaths. Wavelengths are set up between endpoints in the WDM network, avoiding costly electrical–optical (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 the cost per bit) tends to be dominant. Traffic of dissimilar protocols and bit rates needs to be carried, fluctuating heavily in both volume and space. Traditional approaches such as synchronous digital hierarchy/synchronous optical network (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 with 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, refer to [2] and references therein.) OPS-based architectures for metropolitan environments are discussed hereafter.

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

Manuscript received December 15, 2003; revised June 9, 2004. This work was supported in part by the European Commission through the IST project DAVID under Grant IST-1999-11387 and by the Flemish Government through the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT) Generisch BasisOnderzoek and Universiteiten (GBOU) project Optical Networking and Node Architectures. The work of C. Develder has been supported under the Fund for Scientific Research—Flanders (FWO-Vl), Belgium.

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

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

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

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

H. Lønsethagen is with Telenor R&D, N-1331 Fornebu, Norway.

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

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

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

Digital Object Identifier 10.1109/JLT.2004.836773

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 University, Stanford, CA [3], multiple wavelengths are deployed onto a single-fiber ring (in WDM), where each node can receive on only a single (fixed) wavelength, having been 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 multihop scheme requiring excessive E/O conversions, electronic buffering, and packet processing. Switching decisions are made based on packet headers, which are sent using orthogonal frequency-shift-keying (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/time-division-multiplexed (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 paper discusses the MAN architecture devised in the frame of the DAVID (Data And Voice Integration on DWDM) 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 Standard 802.3) or Resilient Packet Rings (RPR, IEEE Standard 802.17). Therefore, this paper will not be restricted to detailing the DAVID architectures and their performance but will 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 DAVID (Data And Voice Integration on DWDM) aims at proposing a viable approach toward OPS by developing networking concepts and technologies for future optical networks [7]. The work ranges from theoretical studies covering traffic studies, control, scheduling algorithms,

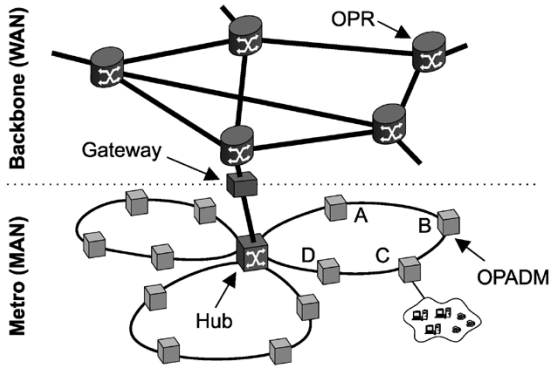

Fig. 1. Generic view of the DAVID network architecture.

medium access control (MAC) protocols, etc., to 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 fibers, each operated in dens-wavelength-division-multiplexing (DWDM) regime (10 Gb/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).

The optical packet add/drop multiplexer (OPADM) ring node puts optical packets [containing client layer traffic, e.g., Internet protocol (IP)] on the ring, using a MAC protocol to decide which time slot at which 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 toward the wide-area network (WAN) through a gateway. This WAN connection from a logical point of view can be seen as an extra ring to and from which to switch traffic. The gateway will be responsible for 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 fiber delay lines (FDLs) to aid in contention resolution [9].

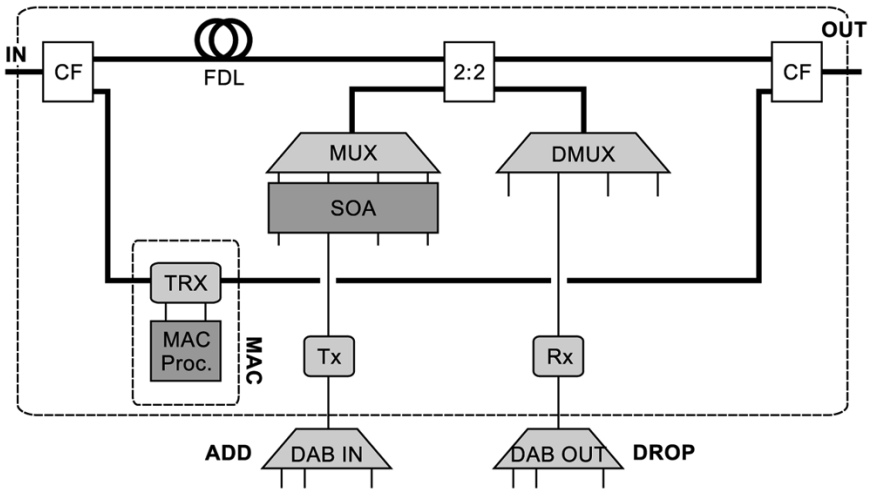

Fig. 2. Passive OPADM node structure. (CF: control channel filter; TRX: MAC transceiver; Proc.: MAC processor; FDL: fiber delay line; SOA: semiconductor optical amplifier array; SOA + Tx: fixed wavelength burst-mode transmitter; Rx: fixed-wavelength burst-mode receiver; DAB: data aggregation board; DAB\IN + DAB\ OUT: aggregation node).

The 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 shown in Fig. 2, the architecture is extremely simple. At the OPADM's input, the control channel first is split off, followed by an FDL for the data channels to account for the delay in processing control information. This processing is achieved by optical–electrical–optical (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 1) to add packets by coupling light coming from a burst-mode transmitter and 2) to drop packets by guiding light to a burst-mode receiver. The WDM channels from multiple receivers and transmitters are separated and combined, respectively, by demultiplexers and multiplexers, respectively. To allow simultaneous add and drop operations within the same time slot, upstream 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.

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, the 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 nonoverlapping connections on the same ring (e.g., from A to B and from C to D in Fig. 1). Moreover, spectral separation of upstream and downstream traffic doubles the amount of required wavelengths.

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 semiconductor 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 reuse (i.e., the dropped slot can be reused for transmission of new data), it does not require separation of upstream and downstream traffic, and a flexible use of the WDM domain (compare 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 Gb/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 therefore, 32 transceivers per OPADM were used in

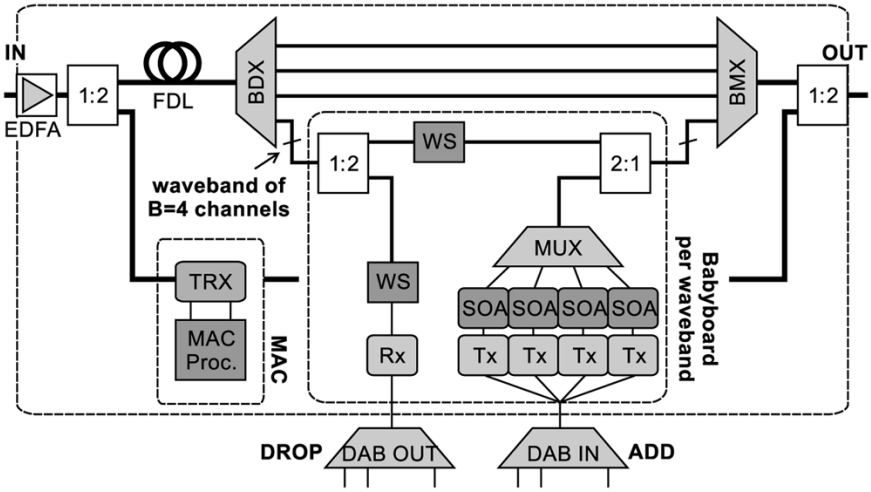

Fig. 3. The active OPADM node structure. (1:2: 1×2 splitter; EDFA: erbium-doped fiber amplifier; FDL: fiber delay line; BDX: waveband demultiplexer; BMX: waveband multiplexer splitting/combining disjoint groups of B wavelengths; WS: wavelength selector for B channels with pass/no-pass gates; $B = 4$; other acronyms: same meaning as in Fig. 2).

the simulations. Second, 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, especially for the passive architecture with its absence of spatial slot reuse, since none of the slots are left over for adding traffic from nodes further upstream. Compared with [10], the physical-layer simulations carried out in this paper are based on components and system parameters that are either identical or very close to the ones measured in the corresponding subsystems 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, for example, 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 baseline 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 (reamplification, reshaping, and retiming) regeneration (thus forming an opaque configuration).

In this paper, 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 fiber (DCF) section was inserted prior to each OPADM to negate the dispersion of the previous single-mode fiber (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

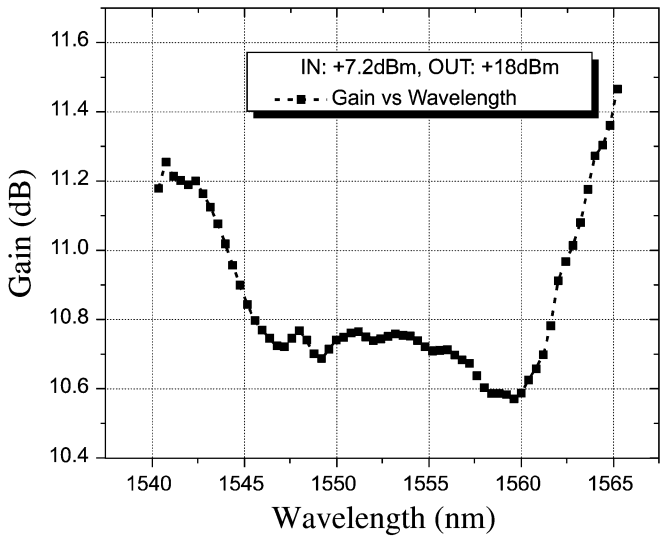

Fig. 4. Gain curve for an EDFA.

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.

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 due to 2R (reamplification and reshaping) regenerative capabilities provided by wavelength converters (WCs) at the OPR's output. This all-optical WC exploits the cross-phase modulation (XPM) of a Mach–Zender interferometer with two SOAs. 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 nonlinear transfer function. The result of implementing a static model is that the chirping effects introduced by the converters (which interfere in

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	2nd 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	

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 cascadability, the Q factor was used

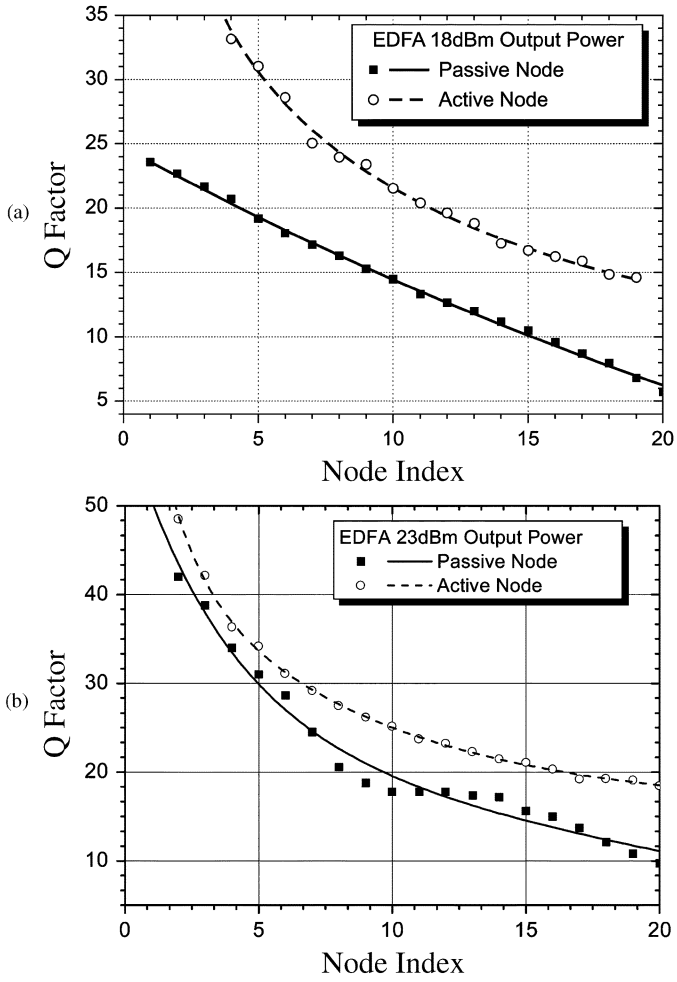

Fig. 5. Q factor versus node index assuming two-stage EDFAs with a 1-dB gain ripple of (a) +18 dBm and (b) +23-dBm total output power.

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 fiber nonlinearities 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. 5(a), 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.) In addition, Fig. 5(a) 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 with 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.

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 seven nodes if an EDFA with +23-dBm 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 the 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 ten OPADMs can be cascaded in both cases. When a lower power EDFA is used, the main limiting factor is the power depletion, while in Fig. 6(b) the small cascadability is caused by fiber nonlinearities.

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.

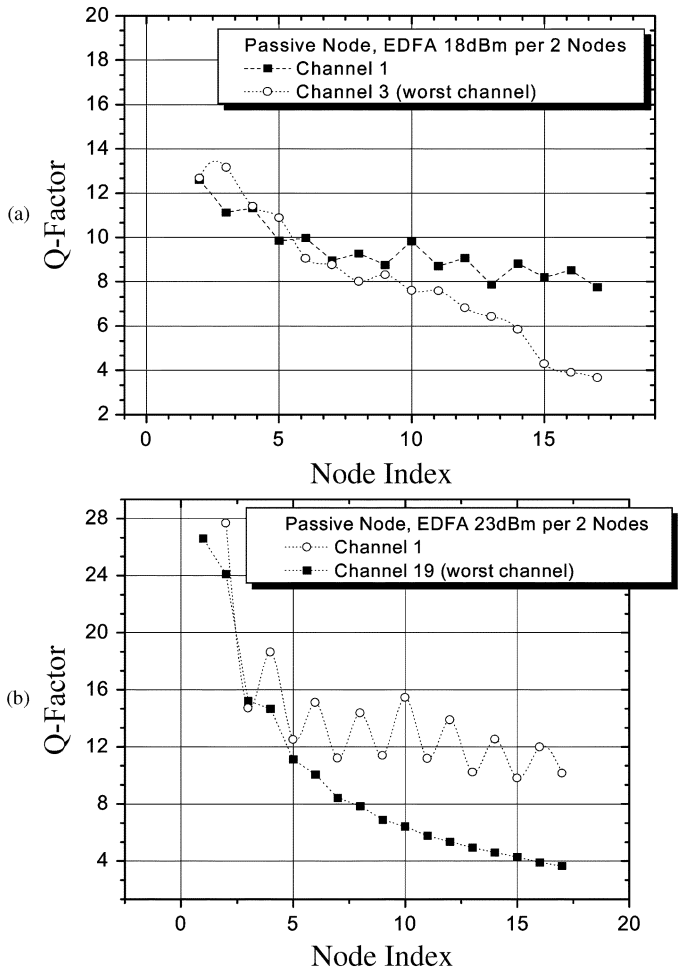

Fig. 6. Q factor for one EDFA per two passive nodes for (a) +18-dBm and (b) +23-dBm amplifiers.

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%

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 the SDH ring, star Ethernet, and RPR. The 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, and jitter). Finally, we used capital expenditure (CAPEX) and operational expenditure (OPEX) models in order to obtain cost benchmarking using the results of the dimensioning studies.

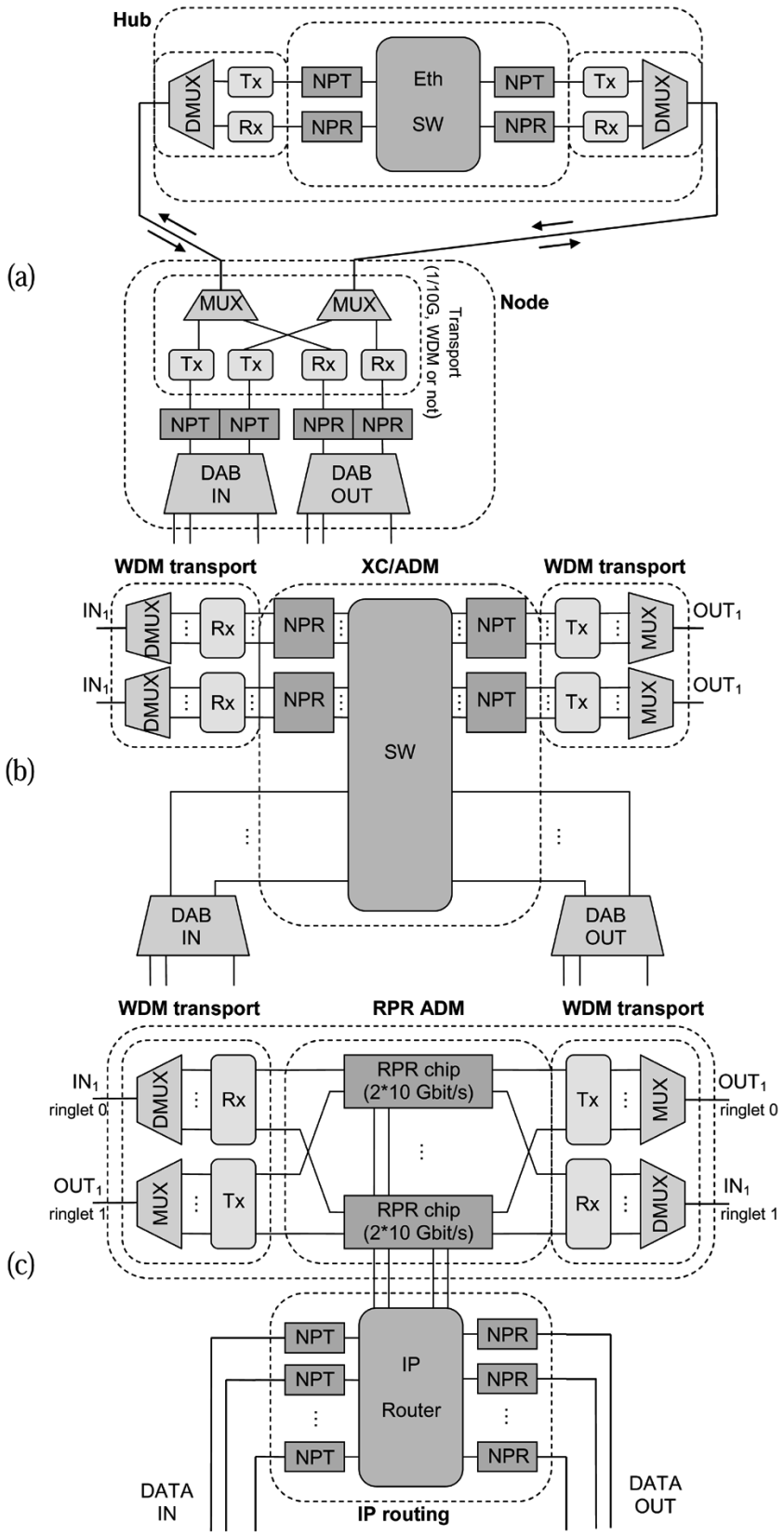

Fig. 7. OPADM node structures. (DMUX: wavelength demultiplexer; MUX: wavelength multiplexer; NPR: network processing receiver; NPT: network processing transmitter; SW: STM-1/STM-4 switch; Eth SW: Ethernet switch; XC: cross connect; other acronyms: same meaning as in Fig. 2). (a) Point-to-point Ethernet Hub + Node. (b) SDH node. (c) RPR node.

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: one server node, two big nodes, four medium nodes, and nine small nodes. We also con-

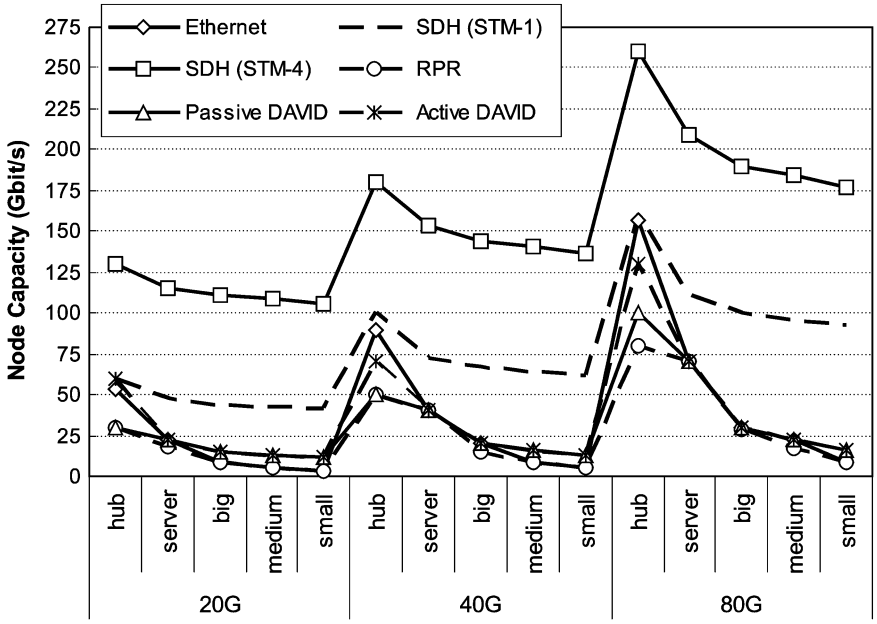

Fig. 8. Node capacity (in gigabits per second) required in the different network architecture for the three traffic volumes.

sidered three different mean traffic volumes: 20 Gb/s (20 G), 40 Gb/s (40 G), and 80 Gb/s (80 G). In addition, we fixed the ratio between the upstream 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 cascadability 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 midterm metro networks. For longer term approaches, we extended the capacity per ring up to 160 or 320 Gb/s while also increasing the number of rings to reach a total capacity close to 1 Tb/s (protected).

To compare the DAVID approaches (shown in Figs. 2 and 3) with the classical Ethernet, RPR, and SDH approaches, we adopted the add/drop multiplexer (ADM) node structures shown in Fig. 7. For the Ethernet solution [Fig. 7(a)], 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. 7(b)] and RPR [Fig. 7(c)] 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 the connection to multiple rings as well as add/drop access. The hub in this case is also an SDH XC (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/multiprotocol label switching (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 Figs. 2 and 3, all node architectures include data aggregation boards (DABs) to aggregate the data traffic coming from or going to the client layer.

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 gigabits per second) 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- or 10-Gb/s channels), and transceivers (either 1 or 10-Gb/s).

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 overdimensioning. 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 due to the flexible design and the optical bypass capability. Nonetheless, the waveband concept (which avoids the need of

TABLE III
TRANSPORT RESOURCES REQUIRED IN THE DIFFERENT ARCHITECTURES

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

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%

a full 32-wavelength selector in each ring node) imposes an overdimensioning 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 of using low-bit-rate interfaces but a drawback at the hub due to the nonshared transport resources and the star topology. For instance, it requires 44 fibers for the 40-G 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.

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 us to obtain lower CAPEX. Indeed, the passive DAVID solution is quite competitive even for low traffic volumes, with it being the second least expensive solution after RPR for the 40-G traffic scenario and the least expensive solution for the 80-G traffic scenario. The Ethernet and SDH solutions are in most cases not highly competitive due to the nonsharing of resources.

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 cannot strongly exploit the optical space reuse 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 (such as DBORN [15]). With respect to the passive DAVID CAPEX, we obtained -26%, -24%, and -22% for the 20-, 40-, and 80-G 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 two, the CAPEX cost is closing to the cost of the passive DAVID solution; multiplying by two, the Ethernet becomes the second least expensive 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 four, the active DAVID solution becomes competitive and close to RPR in the 80-G scenario.

Finally, we extended the initial traffic matrices up to 1-Tb/s scenarios (with 160 G or 320 G on four or two 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 the production of higher volumes for optical components) and a nonlinear 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 an SOA array and multiplexer (as presented in 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.

D. OPEX Comparison

For the OPEX comparison, we adopted a common model where annual costs have been calculated as a percentage of the

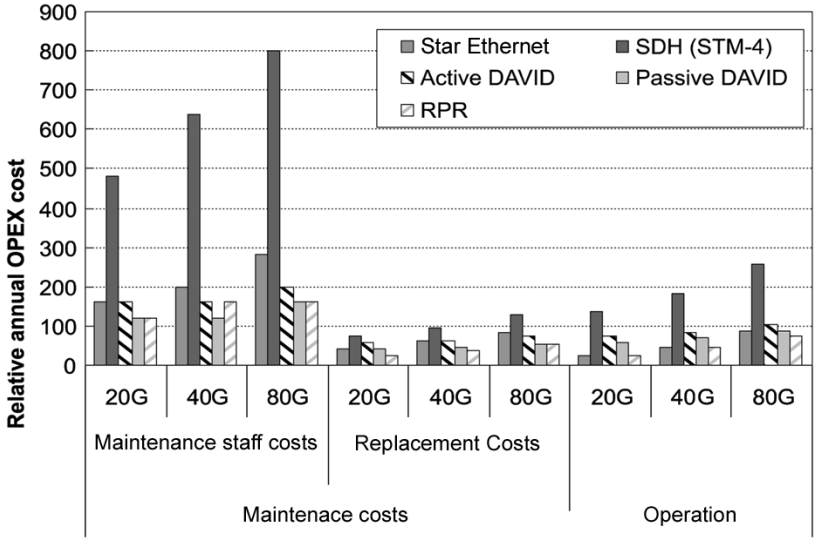

Fig. 9. Relative annual OPEX cost comparison in the different network architecture for the three traffic volumes.

TABLE V
CAPEX COMPARISON

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

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 up 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 differ insignificantly between the various architectures. In addition, migration costs have not been taken into account.

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 labor costs and obviously depends on the required amount of personnel.

The operational costs include all the recurrent costs that 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 fiber $\cdot$ km.

The OPEX costs for the SDH ring are considerably larger than for the other scenarios, since it includes many more network elements, most of which are 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 versus 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.

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.

OPEX sensitivity calculations were carried out to get a better understanding of how OPEX of the different architectures depend on investment costs of various component categories as well as on the mean time between failure (MTBF) of the different component categories. The aim was also to get a better understanding of the OPEX cost model itself. The focus was on

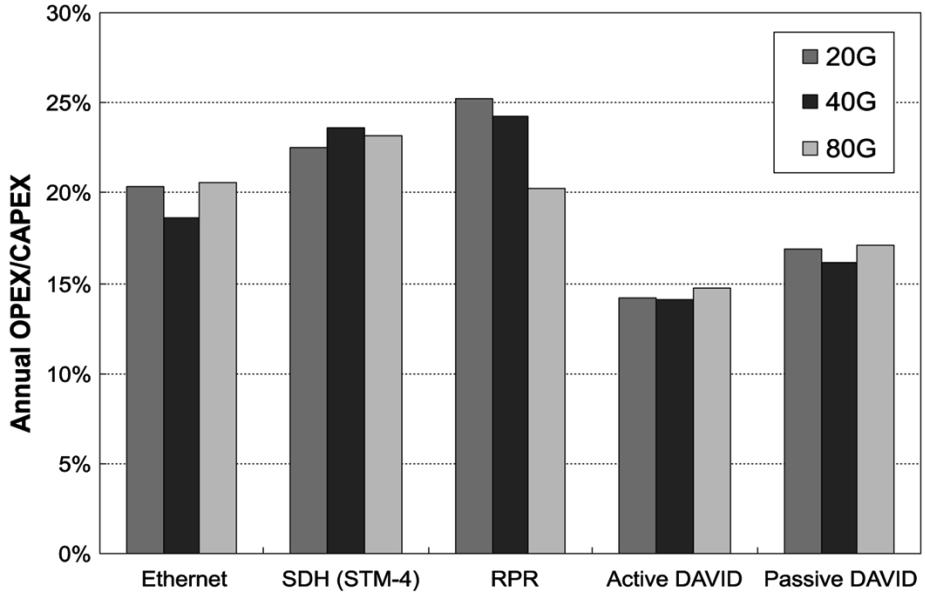

Fig. 10. Annual OPEX over CAPEX.

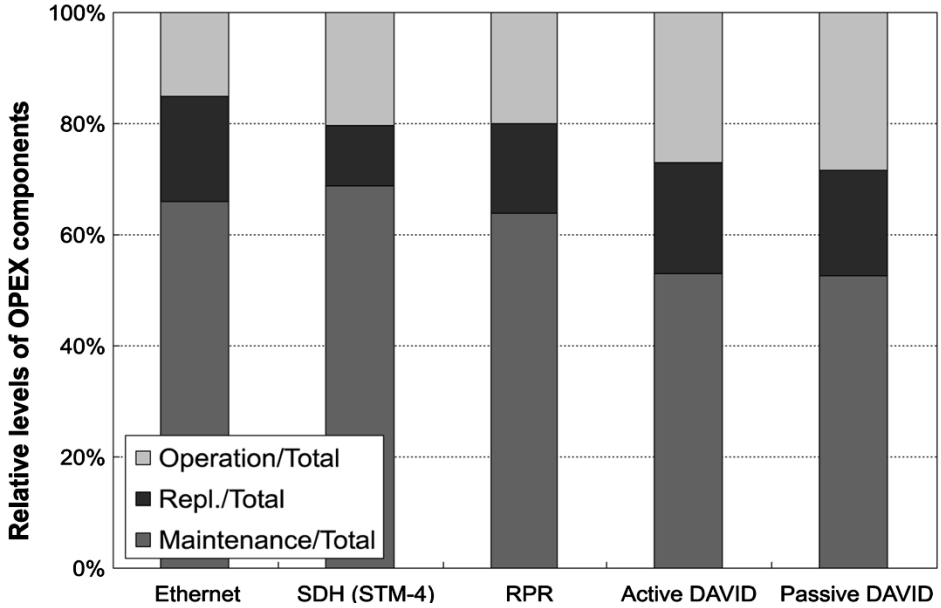

Fig. 11. Relative levels of OPEX components, averaged over the 20-, 40-, and 80-G cases.

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 fiber 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/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 with the traditional metro solutions.

E. Conclusions on Cost Effectiveness

From the extensive benchmarking results—whose results have been summarized previously—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 (a few tens of gigabits per second), two advantageous solutions can be identified: the star Ethernet

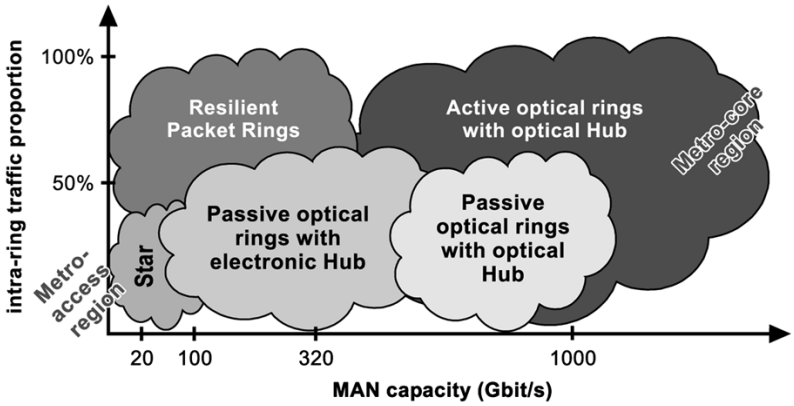

Fig. 12. Possible introduction scenario of the different metro technologies.

(possibly with WDM to share fiber resources) when the ratio of intraring traffic is low, whereas RPR appears the most optimized solution due to space reuse capability.

- At a short/medium term with increasing access bit rate and resulting metro capacity in the range of tens to a few hundreds of gigabits per second, the 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 and gigabit-capable passive optical networks (GPONs)], the capacity in the metro can reach hundreds of gigabits per second to 1 Tb/s. In this case, the two DAVID solutions become competitive due to the optical transparency both at the OPADM and hub levels.

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:

1) transmitter/receiver (Tx/Rx) capacity: the total number of Tx/Rx elements used, summed over all OPADMs;
2) link capacity: the number of wavelengths effectively used per link, summed over all physical links;
3) number of lambdas: the number of wavelengths used per ring, summed over all rings.

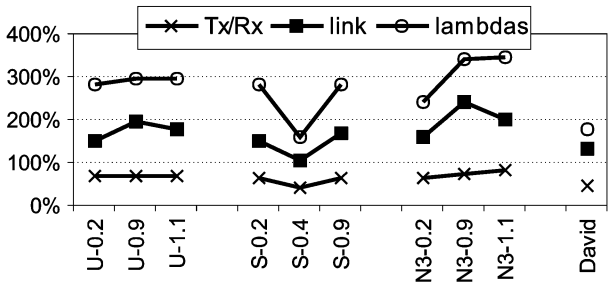

Fig. 13. Cost ratio passive/active; x-axis labels denote demand (U = Uni, S = Serv, and N3 = Neigh3; the number after the dash is the value of d).

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 four demand patterns.

1) *Uni*: This is a uniform demand pattern, where between each two OPADMs a bandwidth d needs to be set up $(D(i,j) = d)$.
2) *Serv*: There is one server node s, which dominates the demand matrix ($D(i,s) = D(s,i) = 2d$, other $D(i,j) = d$).
3) *Neigh3*: Each node only communicates with three other nodes ($D(i,i+1) = D(i,i+2) = D(i,i+3) = d$, the rest is zero).
4) *David*: This is 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 (note that $B = 1$ 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

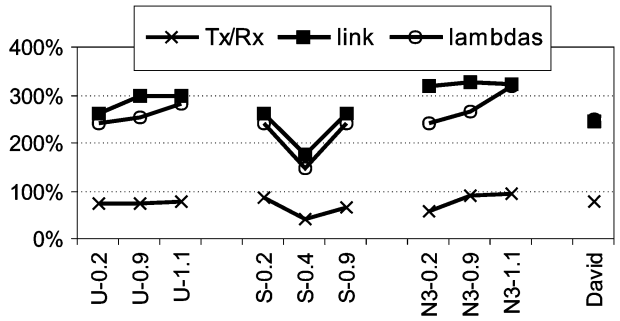

Fig. 14. Cost ratio (bands, $B = 4$)/(no bands, $B = 1$); x-axis labels denote demand.

is no spectral separation for upstream 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 (i.e., 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 OPADM's to add/drop any channel.)

B. MAC Operation and Packet Scheduling

The previously described comparisons between different architectures are based upon a particular static traffic matrix: OPADM and hub dimensioning are 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 (I/O) 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 I/O channels in disjoint pairs through an I/O 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 I/O rings are selected in disjoint pairs, and the number of wavelengths in each ring must be the same to avoid contentions. The I/O 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 OPADMs are notified of the ring-to-ring permutation sequence generated by the scheduling algorithm at the hub 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 to 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 previously—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

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

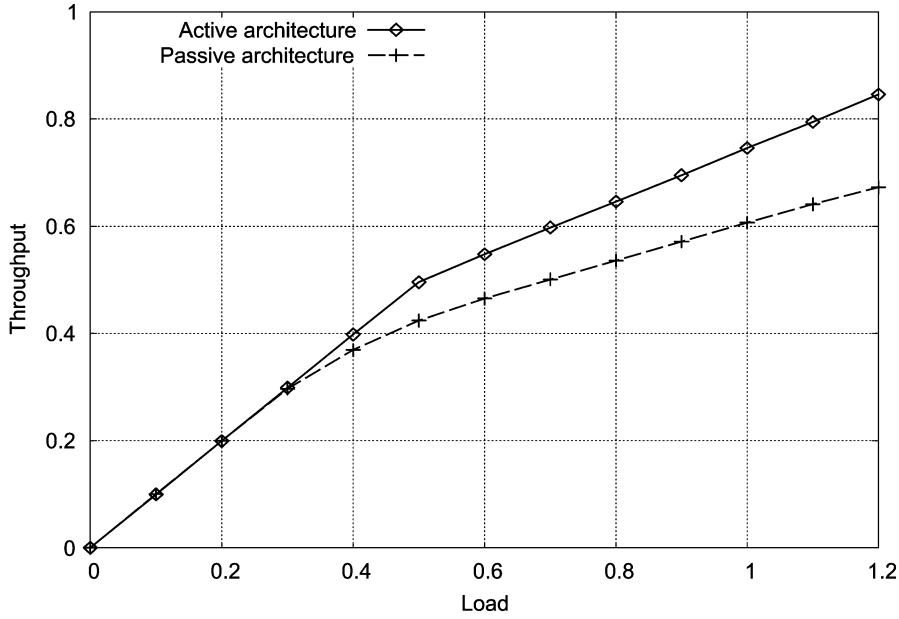

Fig. 15. Overall network throughput under the realistic traffic pattern.

separation of upstream 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 is tunable, is capable of transmitting and receiving only on one channel at a time, i.e., in a particular time slot.)

We consider a network with four 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 multiring 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. Therefore, 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 hereafter, 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.

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 wold 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).

As expected, the passive architecture results in overloaded channels and worse performance due to the nonoptimal 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, nonreported simulation results show that the observed performance degradation is 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 slow reallocation of 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 rethink about what degree of fast tunability/switching is really required in OPS networks in general. Typically quality-of-service (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. Conclusion

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, the authors 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, the assets of active and passive structures were compared 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. It was also shown that exploiting wavelength conversion greatly improves throughput for medium to highly loaded rings.

In conclusion, it has been 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 IST-DAVID project for many fruitful discussions.

References

[1] P. Zhou and O. Yang, "How practical is optical packet switching in core networks?," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM 2003)*, vol. 5, San Francisco, CA, Dec. 1–5, 2003, pp. 2709–2713.

[2] C. Develder, J. Cheyns, E. Van Breusegem, E. Baert, D. Colle, M. Pickavet, and P. Demeester, "Architectures for optical packet and burst switches'' (Invited)," in *Proc. 29th Eur. Conf. Optical Communication/14th Int. Conf. Integrated Optics Optical Fiber Communication (ECOC-IOOC)*, vol. 1, Rimini, Italy, Sept. 21–25, 2003, pp. 100–103.

[3] K. Shrikhande *et al.*, "HORNET: A packet-over-WDM multiple access metropolitan area ring network," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2004–2016, Oct. 2000.

[4] A. Takada and J. H. Park, "Architecture of ultrafast optical packet switching ring network," *J. Lightwave Technol.*, vol. 20, pp. 2306–2315, Dec. 2002.

[5] D. Dey, A. van Bochove, A. Koonen, D. Geuzebroek, and M. Salvador, "FLAMINGO: A packet-switched IP-over-WDM all-optical MAN," in *Proc. Eur. Conf. Optical Communication (ECOC 2001)*, Amsterdam, The Netherlands, Sept. 30–Oct. 4 2001, pp. 480–481.

[6] A. Carena, V. Ferrero, R. Gaudino, V. De Feo, F. Neri, and P. Poggiolini, "RINGO: A demonstator of WDM optical packet network on a ring topology," in *Proc. 6th IFIP Working Conf. Optical Network Design Modeling (ONDM 2002)*, Turin, Italy, Feb. 4–6, 2002, pp. 183–197.

[7] L. Dittman *et al.*, "The European IST project DAVID: A viable approach toward optical packet switching," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 1026–1040, Sept. 2003.

[8] S. Yao, S. J. Ben Yoo, B. Mukherjee, and S. Dixit, "All-optical packet switching for metropolitan area networks: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 39, pp. 142–148, Mar. 2001.

[9] F. Callegati, C. Develder, W. Cerroni, M. Pickavet, G. Corazza, and P. Demeester, "Scheduling algorithms for a slotted packet switch with either fixed or variable length packets," *Photonic Network Communication (PNET)*, vol. 8, no. 2, pp. 163–176, Sept. 2004.

[10] A. Stavdas, S. Sygletos, M. O'Mahoney, H. L. Lee, C. Matrakidis, and A. Dupas, "IST-DAVID: Concept presentation and physical layer modeling of the metropolitan area network," *J. Lightwave Technol.*, vol. 21, pp. 372–383, Feb. 2003.

[11] B. B. Mortensen and M. S. Berger, "Optical packet switched demonstrator," presented at the Int. Conf. Optical Internet Photonics in Switching (COIN-PS 2003), Jeju Island, South Korea, July 2002.

[12] J.-Y. Emery, M. Picq, F. Poingt, F. Gaborit, R. Brenot, M. Renaud, B. Lavigne, and A. Dupas, "Optimized 2R all-optical regenerator with low polarization sensitivity penalty ($<$1 dB) for optical networking applications," in *Proc. Optical Fiber Communication Conf. and Exhibit (OFC 2001)*, vol. 1, Anaheim, CA, pp. MB4-1–MB4-3.

[13] N. S. Bergano, F. W. Kerfoot, and C. R. Davidson, "Margin measurements in optical amplifier systems," *IEEE Photonics Technol. Lett.*, vol. 5, pp. 304–306, May 1993.

[14] F. Davik, M. Yilmaz, S. Gjessing, and N. Uzun, "IEEE 802.17 resilient packet ring tutorial," *IEEE Commun. Mag.*, vol. 42, pp. 112–118, Mar. 2004.

[15] N. Le Sauze *et al.*, "A novel, low cost optical packet metropolitan ring architecture," in *Proc. Eur. Conf. Optical Communication (ECOC 2001)*, vol. 6, Amsterdam, The Netherlands, 30 Sept. 30–Oct. 4 2001, pp. 66–67.

[16] R. Van Caenegem, C. Develder, E. Baert, D. Colle, M. Pickavet, and P. Demeester, "Architectures for OPS metro rings: Comparing active versus passive nodes/A dimensioning point of view," in *Proc. Optical Network Design and Modeling (ONDM 2004), IFIP TC-6 Working Conf.*, Gent, Belgium, Feb. 2–4, 2004, pp. 101–120.

[17] A. Bianco, J. Finochietto, E. Leonardi, F. Marigliano, P. Mitton, F. Neri, and L. Quarello, "Multiclass resource allocation in interconnected WDM Rings," presented at the Optical Network Design and Modeling (ONDM 200), IFIP TC-6 Working Conf., Budapest, Hungary, Feb. 3–5, 2003.

[18] A. Bianco, G. Galante, E. Leonardi, and F. Neri, "Measurement based resource allocation for interconnected WDM rings," *Photonic Network Communication (PNET)*, vol. 5, no. 1, pp. 5–22, Jan. 2003.

C. Develder received the M.Sc. degree in computer science engineering and the Ph.D. degree in electrical engineering from Ghent University, Gent, Belgium, in July 1999 and December 2003, respectively.

He joined the Department of Information Technology (INTEC), Ghent University,in October 1999, working 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 the European Information Society Technologies (IST) projects DAVID and STOLAS, as well as a national research project on Optical Networking and Node Architectures. In January 2004, he left the university to join OPNET Technologies.

A. Stavdas (M'00) 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 wavelength-division-multiplexed (WDM) networks.

Currently, he is Principal Researcher of the Institute of Communication and Computer Systems (ICCS) and head of the Optical Networking Group (http://www.ong.ece.ntua.gr), National Technical Univeristy of Athens (NTUA), Athens, Greece. He is the author or coauthor of more than 50 journal publications and conference articles. He has worked on a number of European projects such as IST-DAVID , IST-OLYMPIC, AC069 COBNET, and AC050 PLANET and also led projects funded by industry (Nortel Networks and Thermophotonics). His current interests include physical-layer modeling of optical networks, optical packet switching, ultrahigh-capacity end-to-end optical networks, optical cross-connect 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 such as IEEE GLOBECOM and Photonics in Switching.

A. Bianco was born in Turin, Italy, in 1962. He received the Dr.Ing. degree in electronics engineering and the Ph.D. degree in telecommunications engineering from the Politecnico di Torino, Turin, Italy, in 1986 and 1993, respectively.

He is an Associate Professor with the Electronics Department of Politecnico di Torino. He has coauthored more than 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.

Dr. Bianco has participated in the technical program committees of several conferences, including IEEE INFOCOM 2000, Quality of Service in Multiservice IP Networks (QoS-IP) 2001, International Federation for Information Processing (IFIP) ONDM (Optical Network Design and Modeling) 2002, 2003, and 2004, and Networking 2002 and 2004. He was technical program Co-Chair of the High Performance Switching and Routing (HPSR) 2003 workshop, and he was involved in several European projects (ACTS-SONATA and IST-DAVID) and Italian projects (EURO, IPPO, and WONDER) on optical networks and switch architectures.

D. Careglio received M.Sc. degree in electrical engineering from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 2000 and the M.Sc. degree in electrical engineering from the Politecnico di Torino, Turin, Italy, in 2001. He is currently working toward the Ph.D. degree at UPC.

He is Assistant Professor at the Department of Computer Architecture of UPC. He is a Member of the Advanced Broadband Communications Centre of UPC (http://www.ccaba.upc.es), and he has recently been involved in European projects ACTS-SONATA, IST-LION, and IST-DAVID. His research interests are in the fields of all-optical networks with emphasis on medium access control protocols, quality-of-service provisioning, and traffic engineering.

H. Lønsethagen received the B.Sc. degree in electrical engineering and computer science from the University of Colorado, Boulder, in 1987. and the M.Sc. degree from the Norwegian Institute of Technology, Trondheim, Norway, in 1988.

He has been with Telenor R&D since 1990, where he has been working with telecom management on particular network-level management of transport networks. From 1994 to 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 and 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 as well as techno-economic analysis.

J. P. Fernández-Palacios Giménez was born in Hellín, Spain, in February 1975. He received the Telecommunications Engineer degree from Polytechnic University of Valencia, Valencia, Spain, in June 2000, where he he conducted his final project on the simulation of wavelength converters.

He joined Telefónica I + D in September 2000, where he has been working on the analysis and evaluation of new technologies either in access or backbone networks. He has participated in Europeans 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, specializing in micro- and optoelectronics from the University of Ghent, Gent, Belgium, in 2003.

She has been working in the Department of Information Technology at the University of Ghent since August 2003.

S. Sygletos received the diploma degree of electrical engineering and computer science from the Faculty of Electrical and Computer Engineering of the National Technical University of Athens, Athens, Greece, in 2000, specializing in telecommunications. He is currently working toward the Ph.D. degree in the field of optical telecommunications at the same university.

He has been involved in the European projects IST-DAVID and IST-OLYMPIC as well as in a number of national research projects. His current research interests include physical-layer modeling of optical networks, all-optical signal regeneration, and fiber transmission issues.

Fabio Neri received the Dr.Ing. and Ph.D. degrees in electrical engineering from Politecnico di Torino, Turin, Italy, in 1981 and 1987, respectively

He is a Full Professor in the Electronics Department, Politecnico di Torino, and leads a research group on optical networks there. 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 coauthored more than 150 papers published in international journals and presented at leading international conferences. He has recently been involved in several European and Italian projects on wavelength-division-multiplexing networks. He is the coordinator of the European 6 FP Network of Excellence "e-Photon/ONe" on optical networks, which involves 38 European research institutions, operators, and manufacturers.

Dr. Neri has served on 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 Modeling (ONDM). He serves on the editorial board of IEEE/ACM TRANSACTIONS ON NETWORKING.

J. Solé-Pareta received the Master's degree in telecommunication engineering and the Ph.D. degree in computer science from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1984 and 1991, repectively.

He joined the Computer Architecture Department of UPC in 1984, where he has been an Associate Professor since 1992. 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 broad-band Internet and high-speed and optical networks with emphasis on traffic engineering, traffic characterization, traffic management, medium access control protocols, and quality-of-service provisioning. He has participated in INFOWIN, MICC, and IMMP ACTS projects, and in IST projects IST-LION, IST-DAVID, and IST-LONG. Within the VI Framework Program, he is participating in NOBEL (an IP project) and in e-Photon/One (Network of Excellence).

M. Pickavet received the M.Sc. and Ph.D. degrees in electrical engineering, specializing in telecommunications, from the University of Ghent, Gent, Belgium, in 1996 and 1999, respectively.

Since 2000, he has been a professor at the University of Ghent, where he is teaching telecommunication networks and algorithm design. His current research interests are related to broad-band communication networks (wavelength-division-multiplexing, Internet protocol, multiprotocol label switching and generalized multiprotocol label switching, optical packet switching, and optical burst switching) and include design, long-term planning, and routing of core and access networks. He focuses on operations research techniques that can be applied to routing and network design. He is currently involved in the European IST projects IST-DAVID (Data and Voice Integration over DWDM), IST-STOLAS (Switching Technologies for Optically Labeled Signals), IST-LASAGNE (All-optical Label Swapping employing Optical Logic Gates) and the Network of Excellence on optical networks ePHOTON/ONe. His work has been published in several international publications, both in journals (e.g., IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS, IEEE COMMUNICATIONS MAGAZINE, *European Transacatioins on Telecommunications*, *Photonic Network Communications*) and in conference proceedings.

N. Le Sauze received the M.Sc. degree in physics from the Université de Bretagne Occidentale, Brest, France, in 1996 and the Engineer Diploma in telecommunications 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 has been leading a study on optical packet metropolitan networks. Through his work, he has developed a strong background in optical packet networks, including optical architecture design, logical performance evaluation, and benchmarking studies. He previously contributed to the French Multiservice Optical Network (ROM) project funded by the Réseau National de Recherches en Télécommunications (RNRT) and European IST-DAVID projects.

P. Demeester (SM'98) is a Professor at the University of Ghent, Gent, Belgium. His current interests are related to broad-band communication networks (Internet protocol, generalized multiprotocol label switching, optical packet and burst switching, access and residential, active, mobile, content distribution networks, and grid) and include network planning, network and service management, telecom software, internetworking, and network protocols for quality-of-service support. He has published more than 250 journal or conference papers in this field.