The Perspective of Optical Packet Switching in IP-Dominant Backbone and Metropolitan Networks

Amaury Jourdan, Dominique Chiaroni, and Emmanuel Dotaro, Alcatel CIT, France
Gert J. Eilenberger, Alcatel SEL, Germany
Francesco Masetti, Alcatel US
Monique Renaud, Opto+, France

ABSTRACT

This article will analyze the rationale and technical solutions for the use of optical packet switching techniques for both backbone and metropolitan applications. It will also provide information on state-of-the-art technologies available for medium-term product development.

INTRODUCTION

The transport network is currently experiencing a doubling of the demand every 8–15 months. In the view of most operators, precedence of data traffic over voice is already a reality, and expected to represent 90 percent within three years. Optimizing the network at all levels (core and metropolitan, Fig. 1) for data applications, and the underlying IP support protocol is definitely the trend in the next generation of routing and switching products prepared by all vendors.

If WDM transmission is clearly expected to meet these capacity requirements, the situation is clearly not settled for layer 1 to layer 3 routing and switching technologies. These technologies, although they progress rapidly, are more or less linked with the well-known Moore’s law, which predicts quite accurately the progress of electronic integration. However, the situation is that the traffic evolution might well outpace this evolution, and therefore, switching may shortly account for the largest part of the complexity and cost of the network. Hence, the efforts to develop a new generation of networking products, putting forward scalability and flexibility as most critical specifications.

Optical switching has more than 10 years of research behind it, with recently increased efforts, fueled by the new attraction around optics in general. If “slow” optical switches are bound to find their way in future layer 1 cross-connect systems, switching and routing are still performed electronically. This article will try to show that the time has come to also consider “fast” optical switching for these applications, in both backbone and metropolitan networks, thanks to unique scalability and modularity features, and the recent maturity of the technology.

PART 1: OPTICAL PACKET SWITCHING IN BACKBONE NETWORKS

In the backbone, the optimization is following three different tracks:

• Higher capacity: Scalable transport solutions, well within the terabit-per-second range, are currently being announced for deployment within two years.
• Increased flexibility: Because of the increased burstiness coming from data traffic specifics, more flexible resource allocation and reallocation schemes are considered as key to operational efficiency in next-generation backbone networks.
• Simplified layer stack: To meet both scalability and flexibility requirements, it seems obvious that routing and switching technologies have to be simplified to alleviate possible bottlenecks, and adapted for the dominant framing format, namely IP.

Of course, this evolution must be compliant with more traditional backbone requirements, such as long-distance transport, high degree of resilience, and effective management. Wavelength-division multiplexing (WDM) transmission techniques are obviously at the heart of any solution, with a deployment roadmap quite easily matching the traffic demand.

Vendors are, however, proposing different alternatives with respect to the layer 3 to layer 1 routing and switching stack, pushing either optical or electronic technologies, IP or other routing approaches.
To keep with the pace of evolution, backbone network elements of multiterabits-per-second capacity will be shortly required, in association with terabit-per-second WDM transmission capacity.

The current trend in development of networking systems for the core of the network is based on two different approaches (the two left protocol stacks of Fig. 2).

**THE IP MULTITERAROUTER APPROACH (T-ROUTERS)**

Based on current technology, Internet Protocol/multiprotocol label switching (IP/MPLS) routers are now reaching capacities in the subterabit-per-second throughput range, with 10 Gb/s line cards. Reaching multiterabits-per-second throughputs requires massive parallelism, and complex interconnection, which, although probably feasible in the medium term as claimed by several vendors for a third generation of router (Fig. 3), are not likely to drive the cost down significantly. However, this approach has the merit of being completely IP-driven, and therefore obviously compatible with traffic engineering and distributed management techniques developed for IP networks.

**THE WAVELENGTH CROSSCONNECT APPROACH**

Introducing wavelength crossconnects, closely associated with edge IP routers: this approach is the most likely to be introduced soon, since it provides scalability at a reasonable cost. It introduces two new features: optical switches (e.g., micromechanical switches, MEMs) for direct crossconnection of optical channels without electrical demultiplexing, and MPLS-based control plane, for enhanced management flexibility. Application of MPLS techniques, allowing switching and transport of IP over service-sensitive lower-layer technologies, on top of optical crossconnects for dynamic wavelength allocation (also called multiprotocol lambda switching, MPLS) provides some flexibility and rapidity in the redistribution of resources along the network, to cope with the massive traffic pattern variations core networks experience over long periods of time.

However, it may not be sufficient in the long run to cope with the large and rapid fluctuations expected from data traffic, even when entering the backbone. Besides, simulations show that the size of the edge routers associated with these crossconnects will still be very large, in the vicinity of 20–30 percent of the crossconnect itself, to provide efficient traffic aggregation, and to take into account the expected multipath traffic due to the lack of efficiency generated by the crossconnect rigidity.

**A COMPARISON OF THE TWO APPROACHES**

Figure 4 illustrates the results of simulations comparing the efficiency of resource utilization in the presence of data traffic using multiterarouters or crossconnects. It is clear that, even in the backbone, significant savings in terms of resource deployment could be obtained by stick-
ing to packet switching techniques, especially when the number of nodes or bit rate of the switched entity increase. This, however, must be traded off with the higher cost of router ports with respect to crossconnects.


THE OPTICAL PACKET SWITCH APPROACH

Optical packet switching (OPS) has the potential to reconcile the optimization of resource utilization and service differentiation from packet switching, with direct application of MPLS techniques, and the scalability of wavelength crossconnects, via the implementation of a third protocol stack solution (Fig. 2, OPS approach), where IP routers are interconnected to a layer 1/2 optical system performing both traffic aggregation and core switching in a much more flexible way than a crossconnect, but at a lower cost than an IP router. This concept is based on burst switching, first proposed in [1]. In this network approach, depicted in Fig. 3, IP packets are concatenated into optical bursts, of larger dimension, at edge nodes, and are then routed as a single entity through the network, within core optical packet routers [2].

The advantage of this approach is to process only one header for multiple IP packets, relaxing drastically the forwarding speed required from core routers, and to scale up their forwarding capability by at least one order of magnitude, well within the multiterabits-per-second range. In addition, with this approach, it becomes possible to consider WDM ports as a single resource (typically 300–600 Gb/s of capacity), and therefore to improve the logical performance and/or decrease the memory requirements with respect to IP routers with single-wavelength processing capabilities. Burst switching also offers the advantage compared to crossconnects of direct compatibility of IP management techniques. It must be stressed that there is much debate going on worldwide about having either a fixed or variable burst frame length. Variable length is more future-proof and adapted to IP traffic, whereas fixed length offers better switching performance and simpler optical implementation.

This burst switching principle is in addition very optics-friendly, since fast optical switching technology requires some specific framing to avoid the loss of any payload data. Optical switching also offers the perspective of large scalability (current research realization have demonstrated up to 10 Tb/s) in a single-stage configuration, avoiding complex interconnection and improving footprint. It is also less sensitive to an increase in line rate, which is very important in view of the oncoming evolution toward 40 Gb/s transmission.

Although the ultimate target could be to reach an all-optical implementation of such a network, and although some other key functions, such as buffering or regeneration, have been demonstrated in the laboratory using optical means (optical fiber delay lines and nonlinear optical elements), these functions could also be realized using electronics, depending on cost/performance trade-offs, and reducing the time to market.

Later we will provide some hints on where does research stand with respect to multiterabits-per-second optical switching systems.

PART 2: OPTICAL PACKET SWITCHING IN METROPOLITAN NETWORKS

Regarding metropolitan area networks, some slightly different optimization parameters are required:

- Flexible upgrade: The pace of bandwidth demand in metro calls for new solutions, much more flexible and scalable than traditional synchronous optical network/synchronous digital hierarchy (SONET/SDH) rings. Scalability is not just reaching huge capacities, but more being able to upgrade smoothly the system during operation with limited initial investment cost. The target capacity for, say, the next three years is typically in the terabit-per-second range for the whole network, at least for the segment between urban access and regional. WDM is obviously entering this market, and is expected to contribute to the network scalability.

- Optimized resource utilization: Data traffic burstiness is obviously higher than in the backbone due to less efficient statistical multiplexing in a network much closer to the access and usually with simpler topologies. Next-generation metropolitan solutions will need to propose more sophisticated bandwidth and resource allocation management schemes, to propose bandwidth-flexible services at affordable costs.

- Cost per transferred bit: Cost is clearly of major importance in metropolitan area networks. Capacity and flexibility obviously have to be traded off with the added cost, although higher utilization of available resources will drive it down.
Transparency: Since a much greater variety of protocols coexist in the metropolitan market, compared to the core, a high level of transparency with respect to these protocols is expected to preserve the past investment of network operators.

Many architectures and solutions are currently being proposed for next-generation metropolitan area networks, with more and more optics, as the technology matures and costs are driven down. Since these networks act as feeders for long distance traffic, and at the same time provide local connectivity, both meshed and hubbed traffic demands have to be considered.

The protocol stack choices of Fig. 2 still apply to metropolitan networks, with two different approaches in the products under development:

• Pure layer 2/3 routing and switching, with no intelligence at layer 1. Traffic aggregation within the optical channels is done directly in an IP router or an asynchronous transfer mode (ATM)-like switch. This approach offers great flexibility, but, as in the backbone, is expected to potentially not be very cost-effective, since all traffic needs to be routed at layer 3 in each node of the network.

• Combined layer 1 and layer 2/3 routing and switching: This approach (referred to as multiservice gateways) is based on layer 1 switching (currently a SONET/SDH-based time-division multiplexed, TDM, core; very soon with some wavelength switching functionality), associated with some more limited processing at layers 2 and 3 for more efficient filling of TDM/WDM circuits. As in the backbone, this approach offers the advantage of lower-cost switching for transit traffic, and potentially a reduction of the optical-to-electronic (O/E) interface cost in the case of wavelength switching (but at the expense of increased rigidity of the network, and a waste of resources).

Again, OPS could bridge the gap between the two other approaches. OPS can offer increased flexibility in the way wavelengths are used, as in approach 1, but with reduced cost, especially due to the reduced number of O/E interfaces (only required for add-drop purposes, full optical transparency being offered to transit traffic), even further than in approach 2 due to the more efficient wavelength utilization.

At the metropolitan level, a rather appealing solution, depicted in Fig. 6, for short-term implementation of an OPS network is to avoid any buffer and regeneration within the optical add-drop nodes:

• Memory is pushed toward the access nodes thanks to the use of a medium access control (MAC) mechanism, already popular in the case of, for example, passive optical networks (PONs). In this approach, electronic buffers hold the information while contention-free transmission is not possible within the network, and packets are emitted only once loss-free propagation is ensured. This obviously puts some limit on the topology complexity and dimension of the network, but a variety of more or less sophisticated protocols are compatible with a metropolitan range (maximum round-trip time 500 μs).

• Regeneration is avoided internally as much as possible due to the limited transmission distance in metropolitan networks. Were it required for optimum performance, a 2R bit-rate- and protocol-transparent solution could be sufficient in only scarce locations, preserving some level of transparency. This feature is key to reducing the number of opto-electronic transceivers, which account...
The legendary longer time to market of optical technology compared to electronics is becoming less and less an argument since the recent explosion worldwide of activity around optical technologies.

for more than half the cost of a conventional system.

• A ring topology, although not optimal when traffic experiences a complex connectivity pattern, is initially preferred, allowing the optical packet add-drop to remain simple (an optical channel only gets through a 2 x 2 switching function), low-cost and with a limited physical impact on the signal quality. Ring interconnection or mesh implementation would require more complex optical systems, and more need for regeneration, but would offer better optimization of deployed resources.

Such a network would have to be managed at two levels:

• At the optical packet level, through the MAC protocol that solves in real time the contention between nodes

• At a higher level, through the application of MPLS techniques to define service differentiated traffic flows and accelerate the allocation of resources

The combined operation of these two management schemes has to be refined to ensure some stability of network behavior.

It is one of the objectives of a recently started European research program (IST Data and Voice Integration over DWDM, DAVID) to design and demonstrate this network concept.

PART 3: THE STATUS OF OPTICAL SWITCHING TECHNOLOGY

This section aims to give a short survey on the availability and maturity of the key technologies required for OPS.

After 10 years of research, such technologies have not yet been widely used in actual products, and it is worth analyzing the most often quoted reasons for this situation:

• The lack of deep and fast optical memories, preventing us from implementing in optics the same router architectures as in electronics

• The poor level of integration, due to some intrinsic limitations and also the limited effort devoted to the subject compared to the mature silicon industry

It is our view that these issues can now be overcome, not only through recent technical breakthroughs, but also through clever network design, making optimal use of optics and electronics wherever they fit best. It is also important to point out that the legendary longer time to market of optical technology compared to electronics is becoming less and less an argument since the recent explosion worldwide of activity around optical technologies.

Let us review the status of the most critical technologies in the path toward OPS, including fast optical switches as well as packet-mode transmitters, receivers, and regenerators, and memories:

• Optical switches: The technology in our view more likely to come to production soon is based on semiconductor optical amplifiers used as optical gates within a passive shuffle network. Current realizations show that it is possible to integrate within one 4-in module several tens of such gates, each capable of gating at least several tens of gigabits per second [3]. These relatively basic elements can also be integrated with passive functions, such as splitters or wavelength multiplexers, to perform very simple wavelength add-drop functions in a metropolitan ring network. Often quoted drawbacks of this technology are a high noise factor and interchannel crosstalk. A careful design of the technology and system levels has been proven to overcome these problems, at up to 1 Tb/s total capacity.

• Packet mode transmission and regeneration functions: this function has been somewhat overlooked in the past at high bit rates, but is essential for actual product development. The key requirements are compatibility with long-distance transmission, tolerance to packet-by-packet phase and amplitude fluctuations, and minimal penalty with respect to synchronous systems. We recently demonstrated a packet mode 10 Gb/s NRZ/RZ receiver that exhibits a very limited penalty when operated in burst mode, typically 6 dB of tolerance to power variations, and very fast clock and phase recovery (< 10 ns). Regeneration can be achieved either using back-to-back burst mode transmitters and receivers, or based on all-optical means. For instance, a cascade of 100 regenerators using semiconductor laser technology for both clock and data recovery has been demonstrated in our laboratory at 10 Gb/s.

• Memory: Until now, the only viable solution for optical buffering was to use switched fiber delay lines. Such optical buffers, when optimized thanks to WDM, can provide several tens of buffer positions to be shared by several tens of optical channels at 10 Gb/s or more. Optical synchronization, required for efficient operation of such buffers, has also been demonstrated using fibers and switches. Although this already allows some practical applications, this is not in the current view at the level of performance required from, say, IP routers, where hundreds of positions are often needed. Therefore, a short-term implementation of an OPS could use optical or electronic memories, for instance in the form of a shared buffer, to save on system cost. Attractive traffic performance (typically 0.8 load tolerance without internal speedup) has been demonstrated through simulations of multiterabit optical routers with a shared buffer architecture using a limited number of fiber delay lines and taking advantage of statistical multiplexing over WDM ports.

These technologies have already been used to build subsystem or system prototype demonstrations. For instance, error-free operation of an OPS (with electronic regeneration interfaces)
compatible with shared buffering and offering up to 10 Tb/s of capacity (1000 ports at 10 Gb/s) with nanosecond-range switching speed and an optimal number of optical gates was recently presented [4]. These technologies can clearly also be downscaled and cost-optimized to be applied to much simpler, one-module, all-optical WDM packet add-drops, lossless and cascadable (typ. 8).

CONCLUSION

Within the first two sections, we present the rationale for the introduction of optical packet switching within terabit backbone and metropolitan transport networks, mainly for improved scalability in the former and reduced cost of opto-electronic interfaces in the latter.

We also sum up the status of research on some key building blocks for actual implementation of these concepts.

It is our opinion, also backed by the growing interest of many companies and research academies worldwide, that the time has come when both traffic and network needs and technology maturity combine to justify, after many years of research, early deployment of optical-packet-switching-based networking products.

ACKNOWLEDGMENTS

We acknowledge partial funding from the European Commission with the IST framework and valuable inputs from our technical partners within the DAVID research project.

REFERENCES


BIOGRAPHIES

AMAURY JOURDAN (Amaury.Jourdan@ms.alcatel.fr) graduated from Ecole Polytechnique, Palaiseau, France, in 1990, and from Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1992. He joined Alcatel Research in 1992, where he has been involved in research activities about all-optical networking and was involved in European projects such as RACE ATMOS, ACTS OPEN and PElican, IST DAVID. He is currently manager of the Photonic Networks Research Unit of Alcatel Research, Marcoussis.

DOMINIQUE CHIARONI graduated in mechanical from the “Institut Universitaire de Technologie, Aix-en-Provence, France; in physics from the Université de Corte, Corte, France; and in optics and microwave from the Institut National des Télécommunications, Evry, France, where he obtained his engineer diploma of telecommunications. He joined Alcatel CIT in 1990 to work in the laboratories of Alcatel Alsthom Recherche. Since 1990, his main research activity has been focused on optical switching. He has been involved in several European projects such as RACE ATMOS, ACTS KEOPS and REPEAT, IST DAVID; he is currently leading a group working on optical packet switching and acting as project leader of a French RNRT project, ROM. He is author and co-author of more than 80 technical papers and 15 patents dealing with optical switching.

EMMANUEL DOTARO [M] received an M.S. degree from the University of Versailles, France, in 1996. During his Ph.D. at the Institut National des Télécommunications, he was involved in several national and European projects (Poloni- um, Galileo). He joined Alcatel Corporate Research Center in 1999, where he is now responsible for optical network planning and design studies. His main interests are IP over optical internetworking issues, including network and system architectures, traffic engineering, and performance evaluation.

FRANCESCO MASETTI [M] received Master’s and Ph.D. degrees in electronic engineering and computer science from the University of Bologna, Italy, on design and implementation of ATM switches, and a Master’s degree in information technology from Polytechnic of Milan, Italy. He has carried out research and coordinated European projects (ATMOS, MEPHISTO) on ultra-high-speed switching nodes and networks, both circuit- and packet-based, implemented with high-speed electrical and optical technologies. He started and then directed the Alcatel Corporate Research Center in Richardson, Texas, and now he has a leadership role on a project to deploy high-speed optical Internet routers. He has authored more than 70 technical papers and 8 patents, and is member of AIE and NMA.

GERT J. EILENBERGER studied communication engineering at the University of Stuttgart where he graduated as a Dipl.-Ing. in 1980. In 1985 he was awarded a Dr.-Ing. degree by the Institute for Telecommunications, University of Stuttgart. He then joined the Alcatel Corporate Research Centre, Stuttgart, where he is responsible for optical systems and networks activities. He has authored many technical papers on electronic and optical broadband telecommunications and holds several patents. Since 1985 he has been involved in various national and European projects on optical communications.

MONTIQUE RENAUD received her engineer diploma and her Doctor 3rd cycle thesis from Institut National des Sciences Appliquées, Lyon, France, in 1983 and 1985, respectively. In 1985 she joined Laboratoires d’Electronique Philips, where her fields of interest were physics and technology of III-V semiconductor microelectronics and integrated optics. In July 1991 she joined Alcatel and has been engaged in research on InP-based photonic switching devices. She is currently leading a group on components for routing and switching. She has been involved in several European projects (RACE and ACTS), and has in particular coordinated the ACTS ACD03 project KEOPS. She has contributed to about 100 papers in international journals and conferences. She has participated in various conference program committees, including OFC, ECIO, and CLEO-E. She is a member of OSA.