Design of Optical Packet Switching Networks

A. Bianco, E. Leonardi, M. Munafò, F. Neri, W. Picco
Dipartimento di Elettronica – Politecnico di Torino
Corso Duca degli Abruzzi 24 – I-10129 Torino – Italy
E-mail: {bianco,leonardi,munafo,neri,picco}@polito.it

Abstract—**The paper considers optical packet switching networks with slotted operation. A general model of network nodes is introduced, based upon a non blocking switching fabric, and re–circulating fiber delay lines to solve contentions. Given the current large bandwidth availability in optical networks, and the projected limitations of electronic switches, a new approach to network design is proposed, aiming at balancing wavelength and buffer allocation taking the number of switch ports as a constraint. Given the problem complexity, a heuristic solution is proposed, using simple queuing theory to model network links. The effectiveness of the new network dimensioning approach is demonstrated by running a simulation program on manually dimensioned topologies, and on topologies dimensioned using the proposed approach.**

I. INTRODUCTION

The technological evolution of telecommunication networks in recent years has shown that the availability of transmission bandwidth has been growing at a faster pace than the switching bandwidth inside network nodes [4]. This means that in the near future electronic switching will become the performance bottleneck of high–capacity networks. The cost of a switching device (be it an IP router, an Ethernet switch, or an ATM switch) grows more than linearly with the data rate on switched channels. Today, data rates of 2.5 Gb/s are quite common, but the technology is ready for 10 Gb/s, and researchers are actively working on 40 Gb/s transmitters and receivers. Further impairments of high–capacity electronic switching devices are the large required power supply, and the large footprint (i.e., the large volume), which increase the infrastructure costs.

The optical technology [5], [6], [7] has the potential for solving some of these limitations, featuring a switching cost that is largely independent from the channel data rate, and smaller power supply and footprint requirements. Thus, while optics today is limited to point–to–point transmission, it will most likely enter the realm of switching soon. Networks implementing switching functions in the optical domain are often called *all–optical networks*. More difficult, and far from current technology, is instead the adoption of optical technologies for network control. We refer in this paper to all–optical networks with electronic control of switching devices.

Current designs of all–optical networks are based on Optical Cross–Connects (OXC), and operated in (fast) circuit switching mode [5]. OXCs are based upon optical switching fabrics; possible technologies for implementing these optical fabrics are Semiconductor Optical Amplifiers (SOA) [8], Micro Electro–Mechanical Systems (MEMS) [9], and Electro–Absorption Modulators (EAM) [10]. A key functionality for all–optical networking is optical 3R (Regeneration, Reshaping, Retiming), and a number of components and systems to implement this functionality has recently been studied in research labs [11]. Optical 3R will enable a decoupling between network design and transmission system engineering, which is fundamental for the success of all–optical networks. An additional useful functionality for all–optical networking is wavelength conversion [12]. Although subsystems capable of implementing these functions are not available today, we assume that they will be available soon.

Packet switching is not well matched with state–of–the–art optical technology, mainly due to the lack of optical memories, and to the lack of optical processing capabilities (i.e., we do not have optical microprocessors). Moreover, optical switching fabrics are typically not capable of being reconfigured in a packet–by–packet fashion (consider that a 100–bytes packet lasts approximately 40 ns at 40 Gb/s). However, packet switching well matches the dominating IP technology, which today is the main support to user applications. Intermediate solutions between fast circuit switching and packet switching, such as Optical Burst Switching (OBS) [13], are currently studied in optical networks. Long-term network architectures must however address packet switching in the optical domain. This paper focuses on optical packet switching, assuming that most limitations of the optical technology will be soon overcome.

II. OPTICAL PACKET SWITCHING

We assume a slotted network operation and fixed–size packets (called cells) lasting one slot. This means that variable–size (IP) packets must be segmented at the boundary of the optical domain, and typically are fed into the all–optical network as trains of cells. Network nodes operate in a synchronous fashion, meaning that input cells are time–aligned, and the switching fabric is reconfigured at slots boundaries, providing a new input/output permutation. Cell alignment at node inputs is a critical function, which we assume to be available.

Fiber *links* interconnecting network nodes are operated in Wavelength Division Multiplexing (WDM): a number of wavelength *channels*[1] is available on each node–to–node link. The WDM channels of each link are space demultiplexed in individual channels at node inputs, and multiplexed on links at node outputs (see Fig. 1). Arrayed Waveguide Gratings (AWG) [14] are a good technology for these mux/demux operations. Before input demultiplexers, cells must be slot-aligned by suitable slot synchronization blocks, so that the switch can inter-

---

[1] *Links* are fibers interconnecting network nodes, as described by the network topology. *Channels* are WDM pipes on links. The number of channels is the outcome of the network design problem.

0-7803-7632-3/02/$17.00 ©2002 IEEE

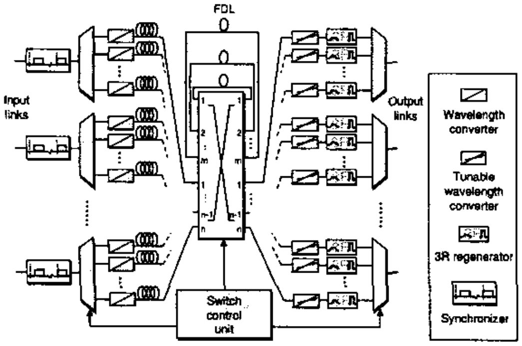

Fig. 1. Architecture of the all–optical packet switching node

nally operate in a slot-synchronous fashion. After input demultiplexers, a wavelength-conversion block is present to adjust wavelengths to switching fabric requirements. Following the input wavelength converters, a Fiber Delay Line (FDL) is inserted to provide the delay required to the switch controller to decide the switching configuration.

The optical switching fabric is non–blocking (any input/output permutation can be configured), and capable of being reconfigured every time slot. Cell buffering capabilities are emulated by means of a set of FDLs [15], [16] connected in feedback mode around the switching fabric. Each of these FDLs provides a delay equal to one slot time, so that a contending cell can be sent to an output connected to a FDL and fed back to the switch input after a delay equal to one time slot. We therefore have n input/output channel ports and m input/output re–circulating ports at each switch. At the beginning of each time slot, an (electronic) control strategy selects packets that are sent to output channel ports among packets available at channel and re–circulating input ports. Priority is given to oldest packets.

Before output multiplexers, optical signals undergo another wavelength conversion to select a unique channel in the output link, and optical 3R.

Note that packet losses are possible in the considered node architecture in case of sustained contention.

III. NETWORK DESIGN PROBLEM

The paper studies the problem of dimensioning a network of optical packet switches (according to the description given in the previous section) interconnected in an arbitrary (mesh) topology.

Given the large bandwidth availability in all–optical networks, the optimization of bandwidth usage, normally considered a must in electronic network designs [17], is no longer a primary network design target. We aim instead at controlling the switching capacity required by the network, i.e., to control the number of switching fabric ports. Note that ports can be devoted to input/output wavelength channels, or to packet re–circulation through FDLs. If the number of ports is fixed, using more ports for FDLs means having a better chance of solving contentions, but also reducing the node–to–node bandwidth, i.e., having less ports for node interconnection, hence less wavelength channels in the network, and a smaller network capacity. This is a new design criterion, which stems from the current scenario in optical networking. A possible formulation of this problem is:

GIVEN: network topology and traffic matrix
FIND: number of WDM channels on each link
OPTIMIZING: network throughput
WITH CONSTRAINT: fixed number of ports at each switch

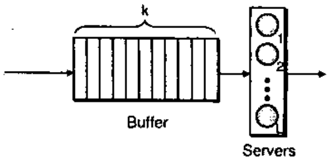

Fig. 2. The link model: an $M/M/K/k$ queue

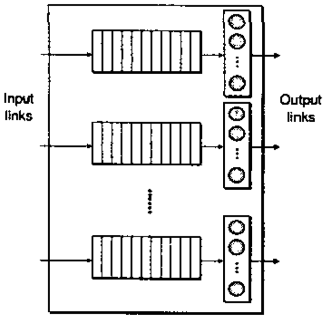

Fig. 3. The node model

Our approach comprises two basis parts:

1. compute the packet loss probability $P_{\text{tot}}(n_1, n_2, \ldots, n_M)$ of the whole network, where n_i is the number of WDM channels on link i; M is the number of links in the network topology;
2. heuristically find an approximation to the minimum of P_{tot}.

Links are modeled with $M/M/L/k$ queues [18], where the L servers represent the L WDM channels on the link, and the queue capacity k refers to the number of FDLs (see Fig. 2 and 3).

The packet loss probability $P_i^{(l)}$ on link i can be therefore computed as follows:

$$P_i^{(l)} = \pi_L \left(\frac{\rho}{L}\right)^k = \pi_0 \frac{\rho^L}{L!} \left(\frac{\rho}{L}\right)^k \qquad (1)$$

where $\rho = \frac{\alpha}{\mu}$ is the offered load, α is the link load, μ is the link capacity, and:

$$\pi_0 = \begin{cases} \left[\sum_{i=0}^{L-1}\frac{\rho^i}{i!} + \frac{\rho^L}{L!}\frac{1-r^{k+1}}{1-r}\right]^{-1} & \text{if } r \neq 1 \\ \left[\sum_{i=0}^{L-1}\frac{\rho^i}{i!} + \frac{\rho^L}{L!}(k+1)\right]^{-1} & \text{if } r = 1 \end{cases} \qquad (2)$$

where $r = \frac{\rho}{L}$.

Note that, by assigning k buffer positions to each queue, we introduce an approximation, since FDLs in the modeled node are instead shared among all output links. Furthermore, our FDLs cannot store a packet indefinitely, since re–circulations degrade the signal quality, so that in our simulations we can impose an upper bound, called MR to the number of re–circulations.

The sequence of packets entering the network at a given node and exiting the network from a given node is called flow. The packet loss probability $P_j^{(f)}$ for flow j is:

$$P_j^{(f)} = 1 - \prod_{i \in L_j} \left(1 - P_i^{(l)}\right) \qquad (3)$$

where L_j is the set of links on which flow j is routed. The network packet loss probability P_{tot} can now be computed as:

$$P_{\text{tot}} = \frac{\sum_{j \in F} t_j P_j^{(f)}}{\sum_{j \in F} t_j} \qquad (4)$$

where t_j is the offered traffic of flow j, and F is the set of flows. This concludes Part 1 of our heuristic procedure.

The mathematical formulation leads to a constrained maximization/minimization of a non linear function in the discrete domain. This is known to fall in the class of NP–hard problems, hence non solvable for non trivial network sizes. We resort, in Part 2, to heuristic approximations of the optimal solution. Given the fixed number of switched ports at each node, we must find the best compromise between channel ports and re–circulation ports. Our approach runs through the following steps.

1. We start from a maximally connected topology $T^{(0)}$, i.e., a topology with the maximum number of channels ports.
2. Topology $T^{(i)}$ is perturbed to find topology $T^{(i+1)}$. To find the perturbation, all links are sequentially considered. On each link, if at least two channels are present (since at least one channel must be present on each network link), one channel port is changed into a re–circulation port (see Fig. 4). The perturbation is found selecting the single port change (from channel port to re–circulation port) that provides the largest reduction in P_{tot}.
3. If $T^{(i+1)}$ exhibits a lower P_{tot}, it is accepted, and we continue from Step 2. Otherwise, the procedure stops.

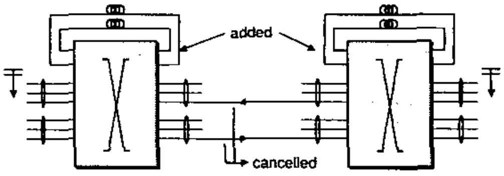

Fig. 4. Topology perturbation

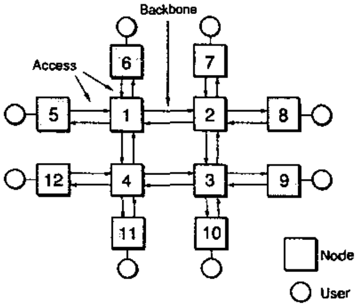

Fig. 5. Regular topology

IV. PERFORMANCE RESULTS

A. Simulation tool

The cell–level CLASS simulator [19] of ATM networks, developed at Politecnico di Torino, was modified to account for the different network setup. The main modification was to allow sets of equivalent parallel channels on the same link. The resulting simulation tool was called SIMON (SIMulator of Optical Networks). The simulation was used to assess the performance of different network designs in terms of cell losses and cell delays. Bernoulli generation of cells was used.

B. Regular topology

The first considered topology is shown in Fig. 5. Four backbone nodes in a regular grid interconnect eight access nodes to which users are attached. Nodes have 16 bidirectional ports to be partitioned between channel ports and re–circulation ports.

Fig. 6 plots packet loss probabilities as a function of the network load in the case of six network dimensioning rules:

1. only 1 wavelength channel is used on each link; the remaining node ports are used for packet re–circulation through FDLs;
2. 2 wavelength channels on each link; the remaining node ports are used for packet re–circulation through FDLs;
3. 3 wavelength channels on each link; the remaining node ports are used for packet re–circulation through FDLs;
4. 4 wavelength channels on each link; the remaining node ports are used for packet re–circulation through FDLs;
5. topology resulting from our optimization technique, based upon $M/M/L/k$ queues; the simulation is run with a maximum of $MR = 4$ packet re–circulations at each node;

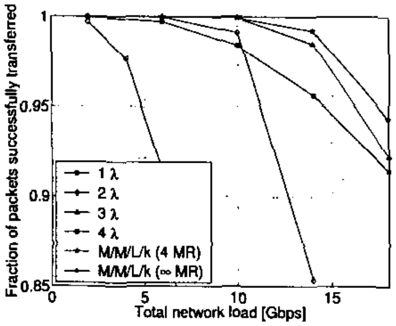

Fig. 6. Loss probability for the regular topology

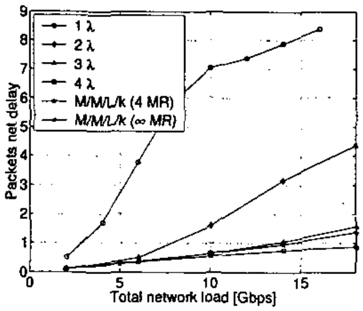

Fig. 7. Delays in the regular topology

6. topology resulting from our optimization technique, based upon $M/M/L/k$ queues, without a maximum number of packet re–circulations at each node in the simulation.

Already for the manual network configurations, with i channels per link, $i = 1, 2, 3, 4$, we observe the expected tradeoff between ports used for contention resolution through FDLs and ports used for node–to–node transmission. Obviously larger values of i lead to lower losses for $i = 1, 2, 3$. Note however that the case $i = 4$, where no FDLs are used in the backbone, leads to larger losses than the case $i = 3$, since the extra bandwidth on the links cannot be used due to packet losses inside nodes. Note also that the optimized design outperforms manual network configurations, thereby confirming the effectiveness of our network design approach.

Fig. 7 shows average packet delays for the same set of network designs. We observe that permitting infinite packet re-circulations does not increase the network throughput, but also does not increase noticeably packet delays with respect to limiting MR to 4 (the corresponding curves overlap in the graph).

C. *"USA" topology*

The second topology for which results are presented is the classical USA backbone [20] depicted in Fig. 8. Again nodes are equipped with 16 bidirectional ports.

Fig. 9 shows packet loss probabilities for network setups similar to those considered for the regular topology. Comments are similar to those to Fig. 6, although we do not observe the performance degradation for increasing number of channels per link. Note that the maximum number of channels is 3, since some nodes have connectivity 5, and number of channels times the connectivity must remain less than the number of ports.

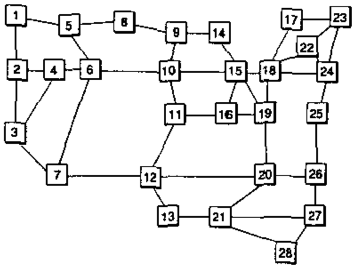

Fig. 8. USA topology

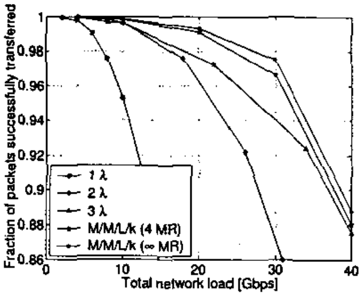

Fig. 9. Loss probability in the USA topology

D. *"Japan" topology*

The last topology for which results are shown is a backbone designed for Japan [20]. See Fig. 10.

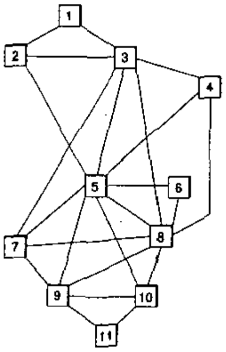

Fig. 10. Japan topology

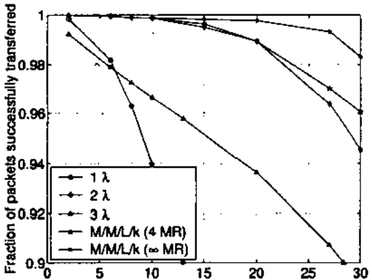

Fig. 11. Loss probability in the Japan topology

Fig. 11 plots packet loss probabilities. Comments are similar to those of previous topologies. We again observe that, for "manual" network configurations, 2 channels per link perform better than 3 channels per link. We also observe in this case a significant performance improvement by allowing an unlimited number of packet re–circulations at nodes, which was only marginal in Fig. 9, and not observable in Fig. 6.

V. CONCLUSIONS

We considered optical packet switching networks with slotted operation. A quite general model of network nodes was introduced, based upon a non blocking switching fabric, and recirculating FDLs to solve contentions. Given the current large bandwidth availability, and the projected limitations of electronic switches, a new network dimensioning criterion was defined, aiming at minimizing the number of switch ports in the network. The optimal formulation of the problem is NP–hard, hence a heuristic solution, using $M/M/L/k$ queues as link model, was proposed. The effectiveness of our network dimensioning approach was demonstrated by running a simulation program on manually dimensioned topologies, and on topologies dimensioned using the proposed approach. Although a number of improvements are possible, and left for future work, the approach proposed in this paper shows that all–optical networks, in addition to new node and protocol architectures, deserve new dimensioning criteria.

ACKNOWLEDGMENTS

This work was supported in part by the Italian research project IPPO (IP Packets over Optics) [1], and by the European Project IST-1999-11742 DAVID (Data And Voice Integration over D-WDM) [2]. Walter Picco worked on this project in the framework of his thesis [3] for the joint M.S. program between Politecnico di Torino and University of Illinois at Chicago.

REFERENCES

[1] `http://ippo.deis.unibo.it/`.
[2] `http://david.com.dtu.dk/`.
[3] W. Picco, *Optical packet switching techniques*, M.S. Thesis, University of Illinois at Chicago, Dec. 2001.
[4] P. Woolnough, "Forecasting traffic growth in the broadband era", Keynote Address at the *5th IFIP Working Conference on Optical Network Design and Modeling*, Vienna, Febr. 2001.
[5] R. Ramaswami, K.N. Sivarajan, *Optical networks: a practical perspective*, Morgan Kaufmann, San Francisco, 1998.
[6] B. Mukherjee, *Optical communication networks*, McGraw–Hill, New York 1997.
[7] T.E. Stern, K. Bala, *Multiwavelength optical networks – a layered approach*, Addison Wesley, Reading, 1999.
[8] J.J. Bernard, M. Renaud, "Semiconductor optical amplifiers", *SPIE OE Magazine*, Sept. 2001, avilable at `http://oemagazine.com/fromTheMagazine/sep01/tutorial.html`.
[9] S.D. Robinson, "MEMS technology micromacines enabling the "all–optical network"", *Proceedings 51st Conference on Electronic Components and Technology*, pp. 423–428, 2001.
[10] D. Marcenac, A. Ellis, D. Moodie, "80 Gbit/s OTDM using electroabsorption modulators", *Electronics Letters*, vol. 34, n. 1, pp. 101–103, Jan. 1998.
[11] B. Sartorious, "3R regeneration for all–optical networks", *3rd International Conference on Transparent Optical Networks*, pp. 333–337, 2001.
[12] J.M.H. Elmirghani, H.T. Mouftah, "All–optical wavelength conversion: technologies and applications in DWDM networks", *IEEE Communications Magazine*, vol. 38, n. 3, March 2000.
[13] S. Verma, H. Chaskar, R. Ravikanth, "Optical burst switching: a viable solution for terabit IP backbone", *IEEE Network*, Dec. 2000.
[14] M.S. Borella et al., "Optical components for WDM lightwave network", *Proceedings of IEEE*, vol. 85, n. 8, pp. 1274–1307, Aug. 1997.
[15] I. Chlamtac, A. Fumagalli, C.J. Suh, "Multi-buffer delay line architectures for efficient contention resolution in optical switching nodes", *IEEE Transactions on Communications*, vol. 48, n. 12, pp. 2089–2098, Dec. 2000.
[16] F. Callegati, "Optical buffers for variable length packets", *IEEE Communication Letters*, vol. 4, n. 9, pp. 292–294, Sept. 2000.
[17] D. Bertsekas, R. Gallager, *Data networks*, Prentice–Hall, 1992.
[18] L. Kleinrock, *Queueing systems – Vol. 1: theory*, John Wiley, 1975.
[19] Politecnico di Torino, dipartimento di elettronica, *CLASS – the user manual*, Ver. 6.20f, May 1998.
[20] K. Murakami, H.S. Kim, "Joint optimization of capacity and flow assignment for self-healing ATM networks", *Proceedings IEEE International Conference on Communications (ICC)*, vol. 1, pp. 216–220, June 1995.