

Efficient Multi-Hop Scheduling Algorithms for All-Optical WDM/TDM Broadcast-and-Select Networks with Arbitrary Transceiver Tuning Latencies

M. Ajmone Marsan, A. Bianco, E. Leonardi, F. Neri, A. Nucci

Dipartimento di Elettronica, Politecnico di Torino

Corso Duca degli Abruzzi 24, 10129 Torino, Italy

Phone: +39 11 564 4000, Fax: +39 11 564 4099

E-mail: {ajmone,bianco,leonardi,neri}@polito.it

Abstract

We focus on optical slotted WDM/TDM broadcast-and-select networks with N wavelengths and N nodes. Nodes are equipped with one tunable transmitter with non-negligible tuning times and one fixed receiver. We propose a novel scheduling algorithm, exploiting the multi-hop approach, to shorten the duration of the transmission scheduling period. A simulation based analysis of the algorithm shows that it leads to significant performance improvements with respect to previous single-hop proposals. We also describe a single-hop scheduling algorithm that performs slightly better than previous proposals

Keywords

Optical Networks, Transmission Scheduling, Multihop Scheduling, Tuning Latency

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35398-2_19](https://doi.org/10.1007/978-0-387-35398-2_19)

G. de Marchis et al. (eds.), *Optical Networks: Design and Modelling*

© IFIP International Federation for Information Processing 1999

INTRODUCTION

All-optical broadcast-and-select networks that use WDM (wavelength division multiplexing) as well as TDM (time division multiplexing) can provide transparent support to several simultaneous very-high-speed end-to-end communications, provided that adequate protocols are defined to dynamically share the available network resources.

A number of access protocol proposals for all-optical WDM/TDM broadcast-and-select networks appeared in the literature, some based on random access approaches (see [1] and references therein), some based on a deterministic schedule of the transmissions from the different traffic sources [2, 3, 4] in circuit switching scenarios; in this paper we discuss novel scheduling protocols of the latter type.

We assume that the all-optical WDM/TDM broadcast-and-select network is based on a broadcast topology (e.g. a star) in which W wavelengths are available for node-to-node communications, and that nodes are equipped with **one** full-duplex transceiver; hence, each node can be source and destination of at most one data flow at any given time. By tuning the nodes transceivers and by dynamically allocating the W available wavelengths, full network connectivity is achieved. Transmission channels are assumed to be slotted and synchronized; each slot can accommodate the transmission of one packet.

The time required to tune the node transceivers is *not* assumed to be negligible with respect to the packet transmission times; in fact, with the presently available optical components, the tuning latency can be equal to several packet transmission times [5, 6]. Access protocols that require a separate tuning action for each packet transmission yield very limited throughput under these conditions. As a consequence, it is necessary to devise simple (hence fast) algorithms to contiguously schedule transmissions on the same wavelength, so that transceiver tuning times do not significantly affect the network throughput.

The problem of finding an optimum transmission schedule in presence of significant tuning latencies was already extensively analyzed [7, 8, 9]. It can be formalized in the following way.

GIVEN a requested traffic matrix R , whose elements r_{sd} are the numbers of packets that must be transmitted from any node s to any node d in a scheduling period, FIND a time/wavelength assignment that guarantees the delivery of the requested traffic, while MINIMIZING the time necessary to accommodate all transmissions, SUBJECT TO technological constraints.

In [7] the single-hop scheduling problem was shown to be NP-hard. Several heuristic approaches for the determination of good schedules were proposed by other authors [8, 9].

Most of the proposals that appeared in the literature aimed at the minimization of the scheduling period duration assuming that all packets must be transmitted in single-hop fashion (i.e., packets are directly transferred from their source to their destination with just one transmission). This may not necessarily be the best approach when tuning latencies are long.

In this paper, we present a novel heuristic approach for the identification of good single-hop schedules, which yields slightly larger throughput than the schedules resulting from previous proposals. Moreover, we show that significant throughput gains can be obtained by transmitting selected packet flows in multi-hop fashion, i.e., by routing them through intermediate nodes. Since we aim at throughput maximization, we do not optimize delays, and we assume that, in the case of multi-hop transmissions, the successive hops that a packet must traverse are scheduled in successive scheduling frames.

The problem of defining an optimal schedule (i.e., a schedule that maximizes the system throughput) when some traffic flows can be transmitted

in multi-hop fashion, will be formalized as an Integer Linear Programming (ILP) problem, hence shown to be NP-hard.

To reduce complexity, we propose in this paper a heuristic approach to the multi-hop scheduling problem solution, based on two separate steps. In the first step, the set of traffic flows that are conveniently transmitted in either single-hop or multi-hop fashion is determined; for each multi-hop flow, the flow routing, i.e., the set of intermediate nodes traversed by the flow, is also defined. As a result of this first step, the single-hop traffic matrix S is derived from R : each element s_{ij} of S is equal to the sum of the number of packets of the traffic flows that, according to the solution found, are to be transmitted in single-hop fashion from node i to node j . In the second step, a single-hop scheduling algorithm is run in order to define the time/wavelength schedule for packet transmissions. The algorithm adopted in this second step represents itself an improvement over previously proposed algorithms.

THE MULTI-HOP SCHEDULING PROBLEM

We consider a network with N nodes, each node being provided with one tunable transmitter and one fixed receiver. We denote by W the number of available wavelengths, and we assume $W = N$, so that each wavelength leads to a different destination. The transmitters tuning latency, i.e. the time required to tune from a wavelength to a different one, is taken to be T slots. Thus, two transmissions from the same source at different wavelengths must be separated by at least T empty slots.

Although we consider in this paper the case of transmitter tunability, both the formalization of the problem, and the proposed heuristic, can easily be extended to the case of receiver tunability, and to a more general context in which both transmitters and receivers are tunable with significant latencies; similar extensions can be devised if the number of wavelengths is not equal to the number of nodes.

Problem Formalization

Define the binary variables v^{ij} that describe the logical topology of the network, i.e. the pattern of logical links in the network, where i and j refer to node indexes. A logical link $i \rightarrow j$ exists if packets are transmitted in single-hop fashion from node i to node j :

$$v^{ij} = \begin{cases} 1 & \text{if a logical link } i \rightarrow j \text{ exists} \\ 0 & \text{otherwise} \end{cases}$$

Define the binary variables b_{ij}^{sd} that describe the routing on the logical topology, where i, j, s , and d all refer to node indexes, such that:

$$b_{ij}^{sd} = \begin{cases} 1 & \text{if the traffic flow } s \Rightarrow d \text{ is routed through link } i \rightarrow j \\ 0 & \text{otherwise} \end{cases}$$

If we assume that the whole traffic flow from node s to node d follows the same route (this constraint guarantees the sequential delivery of transmitted packets to their destinations), the variables v^{ij} and b_{ij}^{sd} must satisfy the following constraints, deriving from such assumption, from the fact that traffic flows must travel over existing logical links, and from the flow conservation at each intermediate node:

$$\sum_i b_{id}^{sd} \leq 1 \quad (1)$$

$$\sum_j b_{sj}^{sd} \leq 1 \quad (2)$$

$$b_{ij}^{sd} \leq v^{ij} \quad (3)$$

$$\sum_i b_{ij}^{sd} = \sum_k b_{jk}^{sd} \quad \text{if } j \neq s, j \neq d \quad (4)$$

Define the binary variables t_l^{ij} , that describe the transmission activity, where superscripts i and j refer to node indexes, and subscript l refers to time slot labels, such that:

$$t_l^{ij} = \begin{cases} 1 & \text{if node } i \text{ transmits to node } j \text{ in time slot } l \\ 0 & \text{otherwise} \end{cases}$$

The variables t_l^{ij} must satisfy the following constraints, deriving from the fact that nodes can transmit and receive no more than one packet in each slot, and from the tuning time constraint:

$$\sum_i t_l^{ij} \leq 1 \quad (5)$$

$$\sum_j t_l^{ij} \leq 1 \quad (6)$$

$$t_l^{ij_1} + t_{l+k}^{ij_2} \leq 1 \quad \forall k = 1, \dots, T \quad j_1 \neq j_2 \quad (7)$$

If all requests in R must be satisfied, the variables t_l^{ij} and b_{jk}^{sd} are related by:

$$\sum_l t_l^{ij} = \sum_s \sum_d r_{sd} b_{ij}^{sd} \quad (8)$$

where r_{sd} , the elements of the request matrix R , represent the number of packets that must be transferred from node s to node d during the considered scheduling period.

Finally, we introduce the binary variables e_l^i to identify an upper bound to the scheduling frame duration. The variables e_l^i , where i is a node index and l the time slot label in a scheduling frame, must meet the following constraints:

$$e_l^i \geq e_{l+1}^i \quad (9)$$

$$e_l^i \geq t_l^{ij} \quad (10)$$

$$e_{l+T-k+1}^i = 1 \quad \text{if } \exists j_1 \neq j_2 \text{ and } \exists k, 1 \leq k \leq T : t_k^{ij_1} + t_l^{ij_2} > 1 \quad (11)$$

Note that inequalities (9) and (10) force $e_i^i = 1$ for all the slots from the beginning of the scheduling period to the end of the last transmission by node i . Inequality (11) forces $e_i^i = 1$ for an additional tuning time, if not enough idle slots are available at the end of the scheduling period or at the beginning of the following one.

With the above definitions, the criterion for the optimization of the multi-hop schedule is the minimization of the transmission period F satisfying:

$$F \geq \sum_l e_l^i \quad \forall i \quad (12)$$

This minimization requires the solution of an ILP problem and, as a consequence, it is NP-hard.

The Heuristic Algorithm

Given the complexity of the multi-hop scheduling problem, we devised a heuristic algorithm which, given a traffic request matrix R , efficiently schedules packet transmissions on an all-optical WDM/TDM broadcast-and-select networks.

Recall that the elements r_{sd} of the traffic request matrix R represent the number of packets that must be transmitted from source s to destination d in a scheduling period. If $L(R)_i$ is the minimum duration of the activity of node i

$$L(R)_i = \max \left(\sum_j r_{ij} + K_i T, \sum_k r_{ki} \right) \quad (13)$$

where K_i is the number of destination nodes to which node i must transmit at least one packet, the length L_{sh} of the single-hop scheduling period (i.e. the period necessary to accommodate all transmissions in R in single-hop fashion) must satisfy:

$$L_{sh} \geq L(R) = \max_i L(R)_i \quad (14)$$

Since we assume that the successive hops of a multi-hop transmission occur in successive scheduling periods, we can split in two separate steps

the problem of defining a packet schedule that minimizes the scheduling period duration, hence maximizes the throughput, while allowing multi-hop transmissions.

The first step aims at the design of a logical topology through which packets must be routed from sources to the destinations; i.e., at the definition of the set of single-hop transmissions as well as the routing for multi-hop transmissions. During the second step, the single-hop scheduling problem is solved starting from a derived traffic matrix S , whose elements s_{ij} are equal to the number of packets that are to be transmitted in single-hop fashion from node i to node j , possibly comprising multi-hop packet flows, according to the solution found in the first step. A novel heuristic single-hop scheduling procedure is proposed, that performs slightly better than previous proposals [8, 9], and that allows the achievement of scheduling period lengths very close to $L(R)$ for non-negligible values of the tuning latency.

We describe the logical topology construction process with an algorithm. Let $S^{(n)}$ represent the single-hop scheduling matrix at step n ; define, for each node i , the length of the transmission schedule at step n .

$$L_i^{(n)} = \sum_j s_{ij}^{(n)} + K_i T \quad (15)$$

The logical topology design algorithm comprises the following steps.

1. Initialize $S^{(0)}$ to R : all connections are enabled (they take place in single-hop fashion).
2. Among all sources of enabled connections, choose node i for which $L_i^{(n)}$ is maximum.
3. Among all destinations of enabled connections whose source node is i , find the destination node j for which $s_{ij}^{(n)}$ is minimum.
4. Find a *pivot* node k so that $s_{ik}^{(n)} > 0$ and $s_{kj}^{(n)} > 0$, through which the flow $i \Rightarrow j$ could be routed in multi-hop fashion; if many such nodes exist, choose the one for which $L(S^{(n)})_k$ is minimum.

5. If a pivot node has been found, compute the temporary matrix S^* as follows: first $S^* = S^{(n)}$, then $s_{ij}^* = 0$, $s_{ik}^* = s_{ik}^{(n)} + s_{ij}^{(n)}$, and $s_{kj}^* = s_{kj}^{(n)} + s_{ij}^{(n)}$.
6. If $L(S^*) < L(S^{(n)})$, then the rerouting of connection (i, j) takes place: $S^{(n+1)} = S^*$.
7. Connection $i \rightarrow j$ is disabled. If any enabled connection is left, go to Step 2; otherwise $S = S^{(n)}$.

In the second step, the single-hop scheduling strategy is solved using the following algorithm.

For each source-destination pair (i, j) , define two functions: $Source(i, j)$ which returns i and $Dest(i, j)$ which returns j . Moreover, let \mathcal{A} denote the set of assigned connections, i.e. the set of source-destination pairs (i, j) corresponding to packet flows whose transmission has already been scheduled; \mathcal{A} is initialized to the empty set \emptyset . For each source node i , define the set $\mathcal{D}(i)$ of destinations corresponding to already scheduled transmissions; $\mathcal{D}(i)$ is initialized to \emptyset for each node i . Let t_s denote the system time and initialize $t_s^{(0)} = 0$. At the algorithm start, all sources and destinations are enabled. The algorithm consists of the following steps.

1. Use a Maximum Weighted Matching (MWM) algorithm (see [10]) over the requests issued by enabled source nodes for transmission toward the enabled destinations, in order to find the switching configuration, i.e. the maximal set of concurrently active transmissions; for each pair (i, j) corresponding to transmissions selected by the MWM algorithm, put $Dest(i, j)$ in $\mathcal{D}(Source(i, j))$, put (i, j) in \mathcal{A} , and disable the corresponding source and destination nodes.
2. Update the traffic matrix S by subtracting all the entries corresponding to transmissions on connections that were just added to

A.

3. For each source node i , evaluate $t_i^{(n)}$, the minimum time at which the node can schedule a transmission toward a destination not belonging to $\mathcal{D}(i)$; let $t^{(n)}$ be the minimum value of $t_i^{(n)}$ greater than t_s . Update $t_s^{(n+1)} = t^{(n)}$. Enable the nodes for which $t_i^{(n)}$ is not greater than $t_s^{(n+1)}$.
4. Enable the available destinations in slot $t_s^{(n+1)}$, i.e. the destinations that are not receiving any packet in slot $t_s^{(n+1)}$.
5. If matrix S is null, the algorithm terminates. Otherwise, consider S_{en} , the submatrix of S referring to enabled source and destination nodes; if S_{en} is not null, go back to step 1. If S_{en} is null, go back to step 3.

Since the algorithm is based on an incremental application of the MWM algorithm, we shall name it IMWM (Incremental MWM).

NUMERICAL RESULTS

In order to provide a first indication about the effectiveness of the proposed multi-hop scheduling approach in all-optical WDM/TDM broadcast-and-select networks, we present some numerical results obtained by running the proposed single-hop and multi-hop scheduling algorithms with an approximate implementation of the MWM algorithm.

We consider networks with a number of nodes N and a number of wavelengths W equal to 32. One hundred different traffic matrices R are randomly generated. The elements of matrix R correspond to individual requests from source nodes, and are taken to be geometrically distributed with mean equal to 10. The elements along the main diagonal are null, meaning that nodes do not transmit to themselves.

The results of the scheduling algorithms are plotted in Figure 1 as curves of scheduling period duration (averaged over the 100 randomly gener-

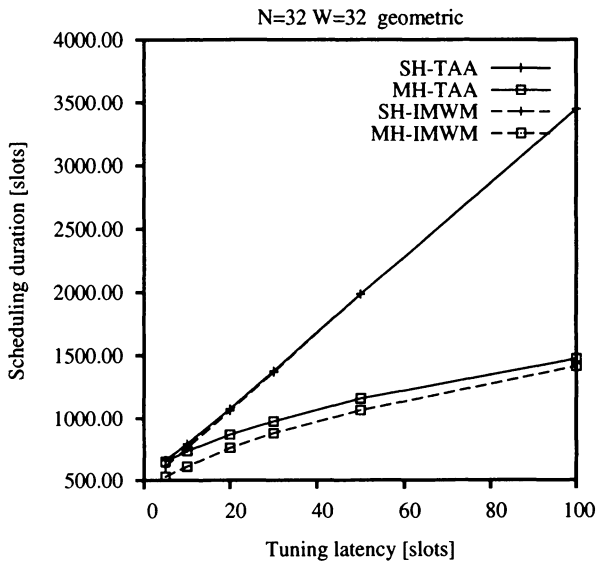


Figure 1: Scheduling period duration as a function of the number of slots necessary to tune node transmitters for the single-hop and multi-hop TAA, and single-hop and multi-hop IMWM algorithms

ated matrices R) versus the number T of slots necessary to tune the transmitters. As regards the single-hop scheduling algorithm, we consider both IMWM, newly proposed in this paper, and TAA, originally proposed in [9]. For the multi-hop scheduling, we always consider our heuristic approach for the design of the logical topology, but again we consider either IMWM or TAA for the second step of the procedure.

The curves in Figure 1 show that the gain obtained with the multi-hop scheduling algorithm can be quite remarkable, specially with long tuning times. Instead, the differences between the two considered single-hop scheduling algorithms do not appear to be significative, even if they slightly favor IMWM over TAA.

Observe that, although not previously mentioned, we can restrict the algorithm to schedule packets with a limited maximum number of hops, aiming at reduced packet delays. In Figure 2 we show the curves of

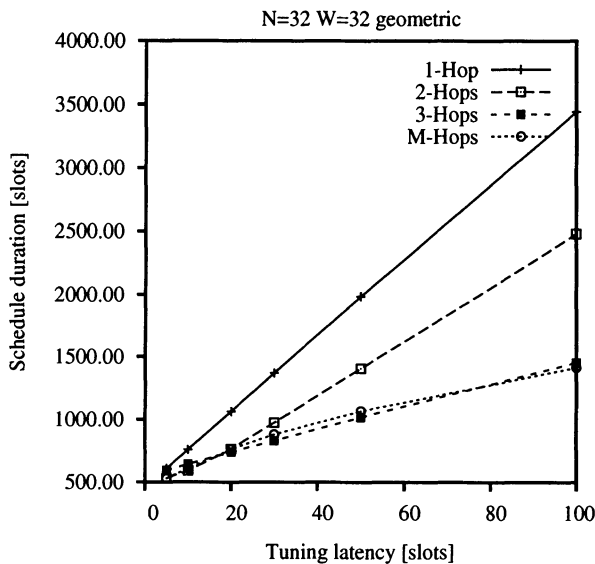


Figure 2: Scheduling period duration as a function of the number of slots necessary to tune node transmitters for the single-hop, 2-hops, 3-hops and multi-hop IMWM algorithms

the average scheduling period duration versus the tuning time in slots, for the case of a maximum number of hops equal to 1 (single-hop), 2, 3, or unlimited (multi-hop). We can observe that the improvements are significant from 1 to 2 hops and from 2 to 3 hops, but there is no evident gain when no restriction on the maximum number of hops is imposed. Actually, the multi-hop case for some tuning delays even produces worse results than the 3-hop case. This is due to the fact that when the number of hops grows, the resulting matrix S is more difficult to schedule in single-hop mode.

CONCLUSIONS

We have considered all-optical WDM/TDM broadcast-and-select networks with N nodes and $W = N$ slotted and synchronized transmission channels. Network nodes have one fixed receiver and one tunable trans-

mitter with non-negligible tuning times.

For these systems, we proposed to adopt the multi-hop scheduling of WDM/ TDM packet transmissions. We first formally defined the multi-hop scheduling problem as an ILP (Integer Linear Programming) problem. To reduce the complexity, we then described a heuristic approach to the problem, which runs through two steps, called the logical topology design and the single-hop scheduling. An algorithm which provides a sub-optimal solution to the logical topology design problem was proposed. We have further introduced an algorithm for the single-hop scheduling problem which provides minor improvements over previously proposed algorithms.

Our results show that the adoption of multi-hop scheduling algorithms provides important throughput advantages over single-hop scheduling. We also showed that limiting the maximum number of intermediate nodes to 2 leads to a good compromise in terms of throughput vs. delay, with respect to single-hop and unbounded multi-hop scheduling algorithms.

REFERENCES

- [1] B.Mukherjee, "WDM Based Local Lightwave Networks – Part I: Single Hop Systems", *IEEE Network Magazine*, vol.6, pp.12-27, May 1992.
- [2] J.C.Lu, L.Kleinrock, "A Wavelength Division Multiple Access Protocol for High-Speed Local Area Networks with Passive Star Topology", *Performance Evaluation Journal*, vol.16, pp.223-239, Nov.1992.
- [3] D.Guo, Y.Yemini, Z.Zhang, "Scalable High Speed Protocols for WDM Star Networks", in *Proc. of IEEE INFOCOM'94*, Toronto, Ontario, Canada, June 1994.

- [4] F.Jia, B.Mukherjee, J.Iness, S.Ojha, "Variable Length Message Scheduling Algorithms for a WDM Based Local Lightwave Network", in Proc. of IEEE INFOCOM'94, Toronto, Ontario, Canada, June 1994.
- [5] A.Sneh, K.M.Johnson, "High-Speed Tunable Liquid Crystal Optical Filter for WDM Systems", in Proc. of IEEE/LEOS'94, Summer Topical Meeting on Optical Networks and their Enabling Technologies, Lake Tahoe, NV, USA, July 1994.
- [6] D.A.Smith et al., "Evolution of Acousto-Optic Wavelength Routing Switch", IEEE/OSA Journal of Lightwave Technology, Special Issue on Multiwavelength Optical Technology and Networks vol.14, n.6, pp.1005-1119, June 1996.
- [7] I.S.Gopal, C.K.Wong, "Minimizing the Number of Switchings in a SS/TDMA System", IEEE Transactions on Communications, vol.COM-33, n.6, pp.497-501, June 1985.
- [8] A.Ganz, Y.Gao, "A Time-Wavelength Assignment Algorithm For a WDM Star Network", in Proc. of IEEE INFOCOM'92, Florence, Italy, May 1992.
- [9] M.S.Borella, B.Mukherjee, "Efficient Scheduling of Nonuniform Packet Traffic in a WDM/TDM Local Lightwave Network with Arbitrary Transceiver Tuning Latencies", IEEE Journal on Selected Areas of Communications, vol.SAC-14, n.5, pp.923-934, June 1996.
- [10] N.McKeown, V.Anantharam, J.Walrand, "Achieving 100% Throughput in an Input-Queued Switch", in Proc. of IEEE INFOCOM'96, San Francisco, CA, USA, March 1996.

AUTHORS' BIOGRAPHIES

Marco Ajmone Marsan is a Full Professor at the Dipartimento di Elettronica of Politecnico di Torino, in Italy. He was born in Torino, Italy, in 1951. He holds a Dr.Ing. degree in Electronic Engineering from Politecnico di Torino, and a Master of Science from the University of California, Los Angeles. Since november 1975 to october 1987 he was at the Dipartimento Di Elettronica of Politecnico di Torino, first as a Researcher, then as an Associate Professor. Since November 1987 to October 1990 he was a Full Professor at the Computer Science Department of the University of Milan, in Italy. During the summers of 1980 and 1981 he was with the Research in Distributed Processing Group, Computer Science Department, UCLA. He has coauthored over 150 journal and conference papers in the areas of Communications and Computer Science, as well as the two books "Performance Models of Multiprocessor Systems" published by the MIT Press, and "Modelling with Generalized Stochastic Petri Nets" published by John Wiley. He received the best paper award at the Third International Conference on Distributed Computing Systems in Miami, Fla., in 1982. His current interests are in the fields of performance evaluation of data communication and computer systems, communication networks and queueing theory. M. Ajmone Marsan is a Senior Member of IEEE.

Andrea Bianco is an Assistant Professor at the Dipartimento di Elettronica of Politecnico di Torino, Italy, since December 1994. He holds a Dr.Ing. degree in Electronics Engineering and a Ph.D. in Telecommunications Engineering both from Politecnico di Torino. His current research interests are in the field of access protocols for all-optical networks and performance analysis of ATM networks.

Emilio Leonardi received the Dr.Ing. and Ph.D. degrees in Electronics Engineering from Politecnico di Torino in 1991 and 1996, respectively. From 1994 to 1995 he was a visiting scholar at the Computer Science De-

partment of the University of California in Los Angeles (UCLA), working on deadlock-free routing and flow control techniques for asynchronous high speed wormhole routing networks. His current research interests are in the fields of performance evaluation of high-speed and all-optical networks, and of queuing theory.

Fabio Neri is an Associate Professor at the Dipartimento di Elettronica of Politecnico di Torino, in Italy. He was born in Novara, Italy, in 1958. He received his Dr.Ing. and Ph.D. degree in Electronics Engineering from Politecnico di Torino in 1981 and 1987, respectively. From 1991 to 1992 he was with the Information Engineering Department at University of Parma, Parma, Italy, as an Associate Professor. From 1982 to 1983 he was a visiting scholar at George Washington University in Washington, DC. In the summer of 1995 he was visiting researcher at the Computer Science Department of the University of California in Los Angeles (UCLA). His research interests are in the fields of performance evaluation of communication networks, high-speed and all-optical networks, discrete event simulation, and queuing theory.

Antonio Nucci was born in Lecce, Italy, in 1974. He currently is working towards his Dr.Ing. degree in Electronics Engineering at Politecnico di Torino. His thesis regards scheduling algorithms and access protocols for optical networks.