ELSEVIER

Computer Networks 32 (2000) 617–631

COMPUTER NETWORKS

www.elsevier.com/locate/comnet

A posteriori versus a priori access strategies in slotted all-optical WDM rings

A. Bianco [a], V. Distefano [a], A. Fumagalli [b], E. Leonardi [a], F. Neri [a],*

[a] *Dipartimento di Elettronica, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino 10129, Italy*
[b] *Center of Advanced Telecommunication Systems and Services, University of Texas at Dallas, Dallas, USA*

Abstract

The paper presents the concept of "a posteriori" access strategy, as opposed to the previously studied concept of "a priori" strategy, to provide a fair and efficient access technique in slotted wavelength division multiplexing (WDM) rings under a large variety of traffic patterns. Using a posteriori strategies, the source node selects the packet for transmission based on the state of the WDM channels in the arriving slot, thus adapting the access decision to the instantaneous traffic in the various channels. The packet selection process of a posteriori strategies, when contrasted with a priori strategies, yields fair and efficient utilization of the ring bandwidth without requiring the equalization of the ring latency, nor imposing restrictions on the traffic patterns allowed in the system. Since the same features cannot be provided by a priori access strategies, a posteriori strategies offers a more versatile access control that justifies the higher implementation complexity due to the multi-channel sensing and on-line packet selection.

Keywords: WDM; Optical rings; MAC protocols

---

1. Introduction

Recent advances in optical technology are opening new business opportunities in the high speed networking field. Interest in the all-optical network concept is re-emerging, both in the LAN/MAN and in the wide area scenarios, as potentially low cost optical devices become available. Indeed, several research and development centers of major telecommunication manufacturers are currently designing all-optical infrastructures for campus and metro networks. All-optical networks are particularly attractive due to their capability of providing access to the huge bandwidth (several THz) available in the fiber without incurring in the so called "electronic bottleneck" due to multiple electro/optical conversions. A practical approach for exploiting the fiber bandwidth is based on wavelength division multiplexing (WDM), that partitions it into a set of channels whose individual transmission rate matches the speed of the electronic interfaces at the user/network interface (UNI). WDM is often combined with either space (i.e., multiple parallel fibers) or time diversity, or both (see [1]).

Depending on the number of available channels and the number of nodes in the network, several system architectures have been proposed in the

---

* Corresponding author. Tel.: +39-011-5644076; fax: +39-011-5644099.

E-mail addresses: bianco@polito.it (A. Bianco), andreaf@utdallas.edu (A. Fumagalli), leonardi@polito.it (E. Leonardi), neri@polito.it (F. Neri).

1389-1286/00/$ - see front matter
PII: S1389-1286(00)00021-9

literature (see [2] for a survey), some dedicating each channel entirely to a single node's transmission (e.g., [3]), some dedicating each channel entirely to a single node's reception (e.g., [4]), and some others providing a number of channels which are dynamically shared by a subset of nodes, or by all nodes, for either transmission or reception purposes (e.g., see [5]).

Multi-channel fiber ring networks have the flexibility to realize any of the above multiplexing architectures [4,6–9] and offer practical survivable techniques against faults [10]. In addition, it is possible to divide time into slots of fixed length, and nodes' transmissions (in the slots) need not be scheduled in advance since both packet collisions on the channels and receiver contention at the destination can be prevented by monitoring the channels prior to the packet transmission [4]. Finally, wavelength stability of tunable transceivers can be achieved with the distribution of reference lines in each channel, and bit and slot synchronization are eased by the linearity of the topology.

The study in [4] showed that the use of fixed receivers and tunable transmitters in the WDM ring is beneficial, in terms of cost and system performance, when compared to other available solutions that provide full connectivity among the nodes, i.e., tunable receivers and fixed transmitters or tunable receivers and tunable transmitters. One of the most important features of this configuration is the fact that with a single fixed receiver at the destination node the cost of the electronic interface is minimized and receiver contention cannot occur. As a result, at the source node the transmitter interface must deal with channel collision only, by means of the access strategy.

The access strategy is part of the medium access control (MAC) protocol and is responsible for selecting the packet, i.e., the channel, for the node's next transmission attempt. Since source nodes that need to transmit to the same destination must share the channel assigned to that destination, the access strategy task is to arbitrate the access to the shared channels in each time slot. The arbitration must be both fair, i.e., nodes see the same network performance, and efficient, i.e., a large portion of the ring bandwidth must be made available to the nodes.

With these objectives in mind, the authors proposed a number of access strategies [4,11–13] that select the next packet transmission using only information that is local to the node, i.e., the destinations of the packets stored in the transmission queues. The advantage of this approach is the simplicity of the electronic control of the UNI as the packet is selected prior to monitoring the channels in the arriving slot. Once the packet is selected, only the channel associated with the packet's destination is monitored by the UNI in order to determine whether or not the packet can be transmitted. If the channel is free the packet is successfully transmitted. If the channel is busy the selected packet is not transmitted in the current slot to avoid collision. Since the packet selection is made by the strategy before the arrival of the next slot, this group of strategies is termed "a priori". In [11–13], a priori strategies which make use of a distinct queue associated with each channel were shown to well exploit the system bandwidth, asymptotically reaching 100% efficiency and guaranteeing a reasonable level of fairness when the distribution of the offered traffic is balanced, i.e., traffic patterns are uniform. Among these access strategies, Synchronous Round Robin (SRR) [11,13] was shown to yield the best performance in terms of achievable throughput and access delay. When the offered load grows evenly in the network, SRR automatically enforces a synchronous scheduling on the nodes' transmissions, thus ideally exploiting the ring bandwidth [13].

If on the one hand the a priori packet selection simplifies the UNI electronic control, on the other hand it limits the range of applications of the slotted WDM ring. In effect, an a priori strategy such as SRR yields fair and efficient access for high loads only if some conditions are met:

- slot framing is used to achieve perfect synchronous scheduling,
- the ring latency is equalized in order to be a multiple of the frame duration,
- traffic is balanced.

As shown in the paper, if only one of these constraints is not met in the ring, the ideal SRR access cannot be maintained, thus causing unfair and inefficient bandwidth utilization. This limitation originates from the non-null probability that the a priori strategy selects for transmission a channel (packet) whose slot is busy while available slots exist on other channels in the same time slot. This is an intrinsic limitation of any a priori access strategy.

The paper presents the concept of "a posteriori" access strategies to provide fair and efficient bandwidth utilization in multi-channel WDM rings under a wide range of system configurations, including absence of frame synchronization and presence of unbalanced traffic patterns. With an a posteriori access strategy the UNI first monitors the whole set of channels to establish which slots on which channels are empty, then it selects for transmission one packet among those whose transmission channel is free. This approach guarantees that under any traffic condition it is always possible to select (if any) a packet that will be successfully transmitted in the arriving slot. In addition, as opposed to SRR, slot framing for synchronous transmission scheduling is not necessary to achieve 100% bandwidth utilization.

It must be observed that the a posteriori strategy requires an increase in both optical and electronic hardware complexity of the UNI when compared to the a priori counterpart. In effect, the a posteriori strategy requires that:

1. all channels are continuously monitored at the source node, thus requiring an array of optical sensors connected to the incoming fiber(s), and an array of (boolean) state variables to select a packet for transmission; note that a priori strategies requires the inspection of only one channel in each slot;
2. the packet to transmit is selected on-line, after detecting the status of every channel, thus requiring fast control electronics, and forcing the serialization of channel inspection and packet selection operations in each slot; conversely, these operations can be pipelined in the case of a priori strategies.

The aim of this study is to quantify the potential advantages offered by a posteriori strategies in terms of both system capacity and level of fairness. A number of experiments comprising various system configurations and traffic patterns are discussed to identify the conditions under which a posteriori access strategies become preferable to a priori access strategies, despite their higher complexity of the source UNI.

2. WDM ring architecture

Without loosing generality, in the remainder of this paper we consider a ring network with a number of channels that equals the number of nodes. Each channel is reserved for the transmissions intended for a distinct destination node; thus a source node willing to transmit a packet to a given destination must use a free slot on the corresponding reserved channel. If we denote by M the number of nodes, M logical channels are thus available in the network. Nodes that need to transmit to the same destination share the same logical channel. Every destination node makes use of one single fixed receiver, and every source node makes use of one fast tunable transmitter. Both optical components are based on proven technology [14]. In addition, receiver contention, i.e., packets transmitted on distinct wavelengths that simultaneously arrive at the same destination node, cannot occur. For more details on the considered node architecture, the reader is referred to [6].

Fig. 1 depicts the logical network architecture for the case $M = 4$. The M logical channels run in

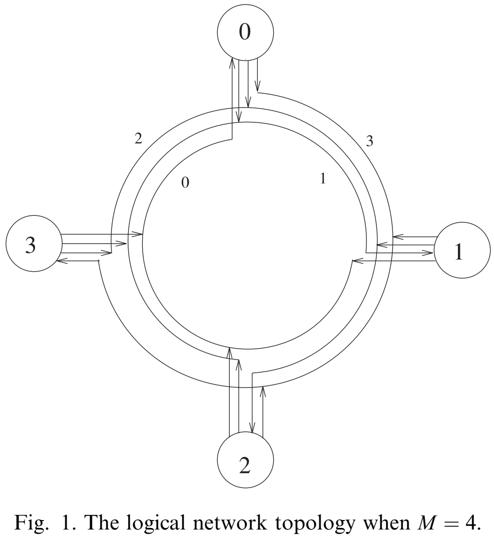

Fig. 1. The logical network topology when $M = 4$.

parallel; they are slotted and synchronized. Full connectivity is achieved by tuning the transmitter at the source node on a packet-by-packet basis. Fixed-length data packets are transmitted, and the slot size is such that one packet transmission time and the transmitter tuning time fit into one slot. Each channel is cleared up at the corresponding destination node, to prevent optical signal re-circulation in the ring. This means that each logical ring is interrupted at a (different) node, and also that the multi-channel ring can be viewed as a set of staggered logical busses folded in a physical ring layout. We assume that nodes do not transmit to themselves.

Slot synchronization in the system is such that slots on different channels are perfectly aligned. In each time slot a node is traversed by a set of slots called *multislot,* i.e., one slot on each channel (which means a slot for each destination), and is free to use one of the empty (unused) slots for transmitting a packet to the corresponding destination node. Only one slot can be used due to the single transmitter assumption. The channel inspection capability required on every channel at each node to avoid collisions among packets directed to the same destination is obtained either by monitoring the optical power in the channels, or by using a subcarrier multiplexing technique (see [15,16]). If contention arise, priority is always given to in-transit packets: as a consequence, once a packet is transmitted, it will reach its destination without collision.

It can be observed that, due to the ring symmetries, each node has a better-than-average access to the channels leading to some destinations, and a worse-than-average access to other channels, leading to some other destinations. If we increasingly number from $i = 0$ to $i = M - 1$ the nodes along the propagation direction of the optical signal in the fiber, the channel on which node i, $i = 0, 1, \ldots, M - 1$, gains the easiest access is the one leading to node [1] $|i - 1|_M$. This occurs because, on the channel associated with node $|i - 1|_M$, node i does not need to defer its transmission to any other node' transmission. Conversely, the channel on which node i gains the most difficult access is the one leading to node $|i + 1|_M$, as, in order to access this channel, node i must defer its transmission to all transmissions of the $M - 2$ upstream nodes. For example, in Fig. 1, node 0 has the easiest access on channel 3 (leading to node 3), and the most difficult access on channel 1, where it must defer to transmissions of nodes 2 and 3. In general, to distinguish among the various transmissions, we define the *access priority* of a packet as the value $|j - i|_M$, where i is the source node and j is the destination node of the packet. According to this notation, 1 and $M - 1$ are the lowest and highest priorities, respectively.

Each node stores the generated packets in $M - 1$ distinct (FIFO) transmission queues, one associated with each destination/channel. This queuing architecture of the source node circumvents the problem of head-of-the-line blocking that may originate in a node using a single FIFO queue [17–19]. Since each queue stores packets with a specific access priority, for simplicity, we refer to each queue using the access priority of the packets stored in the queue.

The presence of multiple queues and one transmitter at the source rises the question: which queue shall the next transmitted packet be chosen from. This decision is made by the MAC strategy described in the next section.

3. Medium access control strategies

The MAC protocol of the described slotted multi-channel ring must arbitrate the nodes' access on each channel to assure that fair and efficient utilization of every channel bandwidth is achieved in the system. Intuitively, the dependence of the access priority on the position of the node along the channel raises fairness issues. An extreme case occurs when a node transmits in all the slots on a given channel, thus starving downstream nodes that compete for the access to the same channel. The key element of the MAC protocol that is responsible for achieving the above objective is the access strategy, adopted by the node to select the next packet for transmission. We recall that the node makes use of a single tunable transmitter

---

[1] The notation $|\cdot|_M$ indicates the modulo M operator.

which allows the transmission of at most one packet per time slot.

We focus our attention on access strategies based on distributed access control that do not require neither slot reservation nor nodes' coordination prior to packet transmission. Some strategies may require to label the transmitted slots with an ordinal number (i.e., a global slot identifier is needed).

The next two sections describe, respectively, the existing access strategies and two newly proposed access strategies.

3.1. Previous work: a priori access strategies

A priori access strategies select the packet to be transmitted in the next arriving slot only on the basis of the distribution of the packets in the transmission queues at the source node, without any knowledge about the channels state. Once the packet has been selected, the channel associated with the packet destination is monitored: if the corresponding slot is free, the packet is successfully transmitted; if the slot is busy, the node refrains from transmission to avoid collision. In each time slot this procedure is executed anew.

A priori strategies are appealing due to their simple control. They do not require continuous inspection of all channels as only one channel must be monitored in each time slot to decide whether or not the selected packet can be (successfully) transmitted. In addition, the packet selection is done off-line, i.e., prior to the arrival of the slot in which transmission will be attempted. The drawback of this approach is that not always the a priori strategy succeeds in selecting a packet whose destination channel is free. Generally speaking, when the selected packet cannot be transmitted due to contention, i.e., the arriving slot is busy on the packet's channel, some system bandwidth is wasted as a different packet could have been chosen by the strategy and successfully transmitted on a different (free) channel. A non-optimal utilization of the network bandwidth may be therefore expected.

The problem of efficient bandwidth utilization was addressed in some previous works by proposing access strategies that reduce the probability to select a packet whose channel is busy [11,12]. Among these efficient a priori strategies, SRR strategy yields an optimal utilization of the system bandwidth under uniform traffic, as shown in [12,13]. This strategy is described next.

3.1.1. The SRR access strategy

The SRR access strategy is based on a cyclic (preferential) scheduling of transmission queues. This scheduling is obtained by first numbering the slots in the ring with an incremental counter s. Note that slots are organized in frames, named SRR frames; each frame comprise $M-1$ slots. In time slot s, node i selects for a tentative transmission the oldest packet with destination $|i+k+1|_M$, with $k=|s|_{M-1}$ (where k is called SRR in-frame-slot-label for s). If the corresponding queue is empty, the transmission of the oldest packet from the longest queue is attempted. If more than one longest queue exists (a tie for the longest queue occurs) SRR selects the packet in the queue with the lowest priority. In any case, if the transmission of the selected packet in not possible because it would generate collision, a new selection is made in the next slot (labeled $s+1$), according to the described algorithm.

For low network loads, SRR behaves almost like a single-queue FIFO strategy. When all queues are always non-empty, its preferential scheduling deterministically orthogonalizes nodes' transmission attempts. More precisely, for network loads larger than or equal to the channel capacity, nodes behave exactly like in a time division multiple access (TDMA): during a SRR frame, whose length is equal to $M-1$ slots, all nodes have exactly one access opportunity for transmissions toward each other node, and they exploit this access opportunity deterministically, thereby avoiding potential conflicts. In this sense SRR is optimal under uniform traffic conditions.

A detailed study of the behavior of the SRR access strategy can be found in [12].

3.2. A posteriori access strategies

A posteriori access strategies rely on the knowledge of both the distribution of the packets in the transmission queues of the node and the current status of the channels. First, the node monitors the channels and detects those with an empty slot. After this preliminary detection, a packet (if any) is chosen from the queues associated with the free channels and successfully transmitted. We say that non-empty transmission queues corresponding to destinations whose arriving slots are sensed free are *enabled* for transmission. An a posteriori strategy thus selects one of the head-of-the-line packets in the enabled queues.

With an a posteriori strategy a node is always able to transmit a packet if it has at least one enabled transmission queue. Thus, the bandwidth efficiency issue of the access protocol is taken care of at the cost of some additional complexity at the UNI, as opposed to the approach provided by the a priori strategy SRR that requires numbered slots.

In the optical layer this complexity consists of an array of sensors that continuously monitor all the channels. At the electronic level this complexity implies that the node must process the information derived from every channel in parallel, thus, requiring additional fast electronics when compared to the a priori approach. In terms of processing, the selection of the packet must be done on-line, *after* the state of all channels has been detected, i.e., the detection and selection operations must be serialized, a problem that does not arise in a priori strategies where the selection of the packet and the channel inspection can run in parallel. This serialization leads to an increased latency for in-transit traffic at each node.

Although based on a straightforward concept, a posteriori strategies must pay particular attention to the fairness issue. Contrary to a priori strategies where the selection of the packet is totally independent from the status of the channels, in a posteriori strategies the selection of the packet is affected by the traffic originated by upstream nodes. Intuitively, this mechanism tends to concentrate the packet selection on those channels that are more often free. As a result, packets with high access priorities ($M-1, M-2, \ldots$) are more likely to be transmitted than others, thus exacerbating the potential unfairness of slotted ring network. In addition to that, we must note that the node does not have any information regarding the transmission activity of downstream nodes. As a consequence, each node can statistically improve its own packet transmissions in subsequent slots, for example by maximizing the number of transmission queues that are left non-empty after each transmission, but it cannot maximize the transmission probabilities of the other (downstream) nodes.

Two a posteriori access strategies are considered: the sequential priority choice (SPC) and the longest queue choice (LQC). SPC is a direct extension of SRR to the a posteriori case, and like SRR, it performs very well under uniform traffic conditions. LQC instead privileges transmissions on the more loaded channel, with the aim of balancing the occupancy of the transmission queues, thereby improving the overall network performance.

3.2.1. The SPC access strategy

According to the SPC strategy, and similarly to the SRR strategy, node i selects for transmission in slot s the first (oldest) packet in the queue associated with destination $|i+k+1|_M$, with $k=|s|_{M-1}$, under the condition that the queue is enabled, i.e., it is non-empty and the corresponding channel is free. As an alternative, SPC selects the oldest packet in the longest enabled queue. If two or more queues satisfy the condition (a tie for the longest queue occurs) the lowest priority among the longest enabled queues is selected.

3.2.2. The LQC access strategy

Nodes following the LQC access strategy always select for transmission the packet at the head of the longest enabled queue, i.e., among queues storing packets directed to destinations whose channel is sensed free. If more than one longest enabled queue exist (a tie for the longest queue occurs), the lowest priority among longest enabled queues is selected.

The rationale behind the LQC design is the attempt to maximize the number of future transmissions by maximizing the number of queues left non-empty after each transmission, thereby improving overall performance.

3.3. Complexity considerations

With respect to SPC, LQC is simpler to implement, since it does not require any frame synchronization, and queues can be easily linked in a list ordered by number of waiting packets. Note that packet arrivals at each input port in a time slot at most swap two queues in this ordered list, if we suppose that no more than one packet is generated at each input interface in a time slot; thus the complexity of searching for the longest queue does not depend on the number of queues. Furthermore, results presented later in the paper show that LQC is performing better, at least in stationary traffic conditions. LQC can however lead to starvation phenomena: for example, a single packet may remain in its queue forever if it is competing with an overloaded queue. SPC is instead less subject to these undesired behaviors. A strategy similar to LQC, but based on packet ages, would do better, but it would be more difficult to implement, since every packet must be time-stamped, time stamps must be coded with a fixed number of bits, and a search for the oldest packet among queue heads must be performed. Thus, both SPC and LQC represent a good compromise between performance and implementation cost.

For the aforementioned reasons, both SPC and LQC are presented and studied in the paper.

3.4. Equivalence with input-buffer switches

It can be noted that a collision and contention free MAC protocol for our network, in which a channel is provided for each destination, solves (sub-optimally) a problem equivalent to finding a non-conflicting transmission pattern in a non-blocking input-buffered packet switch.

Indeed, at most one packet per slot time can leave each input, and enter each output. Possible contentions are solved by buffering data at the input, and no buffering is considered at the output, since we assume that the receiver can process all incoming information. Given this description of the interface between user transceivers and the network, our multi-ring WDM network becomes equivalent to a *distributed* $M \times M$ input-buffered switch. The solution of contentions for the shared channels can be formally described as the problem of finding a maximal matching in the bipartite graph where one set of nodes represents input ports, the other set of nodes represents output ports, and edges represent the presence of packets to be transferred from an input to an output port. Weights on edges can be used to refer to the number of packets waiting in the corresponding input queue. Scheduling algorithms for input-buffered switches [19,20] provide optimal or approximate solutions to this matching problem. The optimal solution aims at the maximization of either the number of transferred packets (maximum size matching (MSM)), or of the total weight of transferred packets (maximum weight matching (MWM)). Algorithms are known to optimally solve these problems, with computational complexities $O(M^{2.5})$ for MSM and $O(M^3 \log M)$ for MWM [21,22].

Several heuristics were proposed in the literature to solve the matching problem with a lower complexity (see [20,23] for a comparison of different approaches). Complexity can be reduced to $O(M^2)$ with minor performance penalties. This complexity is partly motivated by the fact that, in order to avoid limitations due to head-of-the-line blocking, virtual output queueing [19] is normally required, i.e., at each input a separate queue is maintained for each destination, for a total of M^2 input queues.

The scheduling algorithm in input-buffered switches can be executed either by a centralized controller, or at each input port in a distributed fashion. A major drawback of input-buffered systems in either case is that state information must be distributed, usually wasting transmission resources in the form of signalling channels.

The WDM rings considered in this paper suggest (see also [24]) a way to implement an input-buffered switch in which state information is implicitly distributed, so that no transmission resources are devoted to signalling. The scheduling algorithm becomes what we call medium access protocol, and it is executed in a fully distributed fashion. Complexity at each input port for SRR is $O(1)$ if queues are maintained sorted by queue length, and $O(M)$ if the longest queue must be sequentially searched, while it is $O(M)$ for both SPC and LQC, due to the search among enabled queues. The overall complexity is $O(M^2)$ for a posteriori strategies, and it can be reduced to $O(M)$ for a priori strategies. Thus, complexity is comparable with the simplest heuristics that were proposed to solve the matching problem. For what regards performance, the access strategies for WDM rings are comparable with scheduling heuristics in stationary and uniform traffic conditions, while they suffer from intrinsic topological unfairness in strongly unbalanced and time-varying traffic conditions.

As already mentioned, an advantage of a priori strategies is to permit parallel implementations (pipelining of operations) at each input port, while a posteriori strategies require a sequential execution.

4. Performance comparison

Performance comparison among the three access strategies is performed via simulation. Analytical models for the considered strategies are beyond the scope of this paper. The interested reader is referred to [13] and [25] for models of SRR, while [12] derives analytical results for WDM rings in general. In the simulation model, packets whose fixed length fits into one slot, are generated according to a Poisson process whose rate is denoted by λ_i at node i. The total network traffic is thus $\Lambda = \sum_{i=0}^{M-1} \lambda_i$. The probability that a packet originated at node i is directed to node j is indicated by p_{ij}, with $p_{ii} = 0$. A network with 16 nodes, i.e., $M = 16$, is considered in the experiments reported in this paper. The length of the multi-ring is equivalent to 30 slots (i.e., 6 km at 1 Gb/s for packets of 1000 bits). We assume that each transmission queue has a finite capacity of 100 packets. Packets that cannot find room in the transmission queue are dropped.

4.1. Uniform traffic scenario

Fig. 2 shows the average queuing delay, measured in time slots versus packet priority in a uniform traffic scenario, i.e., $\lambda_i = \Lambda/M\ \forall i$ and $p_{ij} = 1/(M-1)\ \forall i \neq j$. Delays are averaged over all packets with the same access priority at all

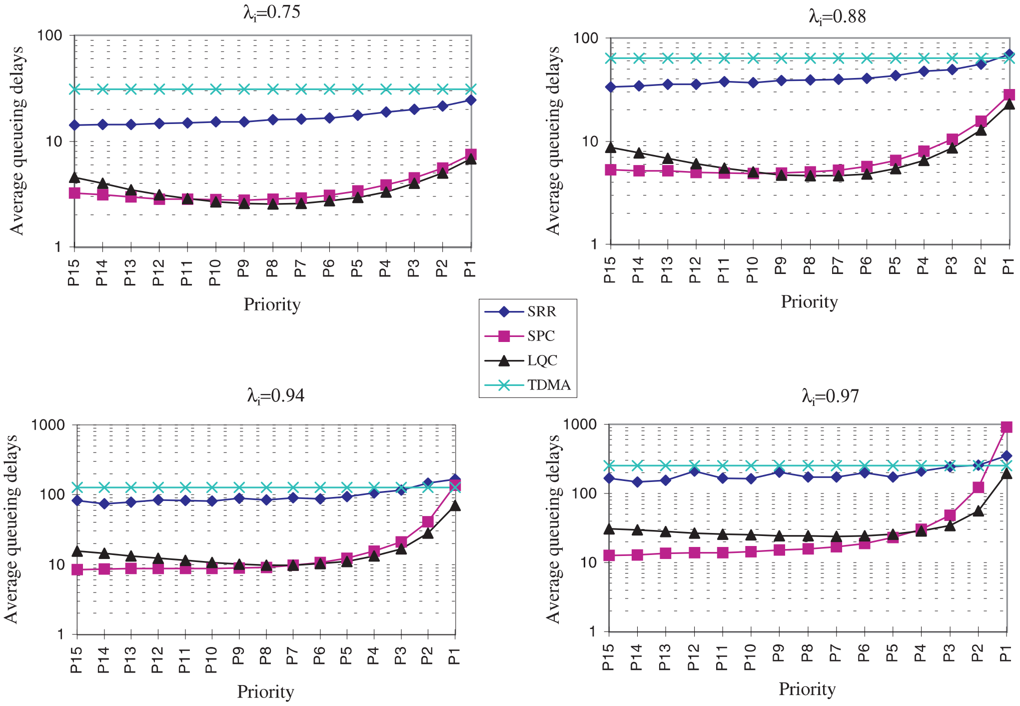

Fig. 2. Queuing delays by priorities for the three access strategies and the TDMA scheme for variable network loads.

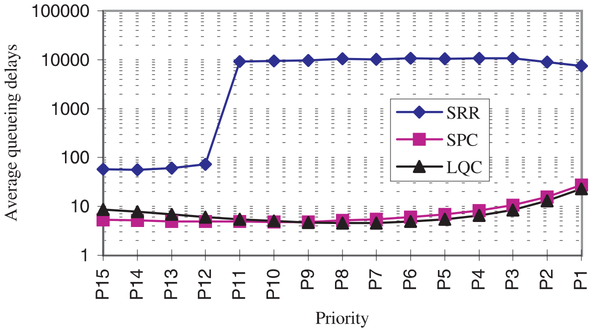

Fig. 3. Queuing delays at node 0 when the ring length is equal to 34 slots.

nodes. Four cases are shown, by varying the offered load: $\lambda_i = 0.75, 0.88, 0.94$, and 0.97.

In addition to the three access strategies, the delays obtained by TDMA, where successive slots are statically devoted to transmissions to successive destinations, are shown for comparison.

It can be seen that SRR asymptotically approaches TDMA for increasing loads. A reduced queueing delay is achieved by the a posteriori strategies. At heavy loads, the gain achieved with the a posteriori strategies decreases since SRR is designed to be optimal in sustained overload.

Focusing our attention on the two a posteriori strategies, it can be noticed that LQC has the tendency to reduce the worst queuing delay (corresponding to the packets with the lowest priority) at the expenses of the queuing delay of the highest priority packets. This result is confirmed by the intuition that, by giving priority to the transmission of packets in the longest queues, LQC attempts to equalize the delay in the transmission queues. In addition, LQC achieves a slightly lower overall average queuing delay since its balancing effect on the average leads to a larger number of non-empty queues, which in turn increases the probability to find a packet for transmission at any given time slot.

4.2. Varying the ring latency

The performance of SRR is sensitive to the ring length. This fact can be explained by considering that, in order to optimally apply the SRR packet selection strategy and to maintain the same SRR in-frame-slot-label during the successive tours of the multi-ring, slot s after one round trip in the ring must return to the node as slot $s + k(M - 1)$, where k is an integer. Only under this condition the source node maintains a perfect Round Robin schedule over the other (destination) nodes. When this condition is not satisfied by the ring length, the TDMA-like orthogonalization of the SRR scheme under heavy load conditions cannot be achieved. As a consequence, the ring length, measured in slot times, must be an integer multiple of the SRR frame, (i.e., of $M - 1$ slots): in Fig. 2 it was taken to be equal to $2 \times (M - 1) = 30$ slots. When the ring length does not satisfy the above constraint, a strong performance degradation is observed.

Fig. 3 reports the queuing delay measured at node 0 versus packet priority for each of the three strategies. Traffic is uniform, with $\lambda_i = 0.88$ and the round trip delay of the ring is chosen to be different from an integer multiple of $M - 1$ time slots. As expected, SRR reveals to be unfair with a delay factor of 100 between the lowest and the highest priority packets. Conversely, despite its synchronous scheduling, SPC achieves a queuing delay that is only marginally sensitive to the variation of the ring length. At higher loads SPC originates a more pronounced unfairness behavior, whereas LQC, due to its asynchronous access, is totally unaffected by the ring length at any load.

4.3. Client–server scenario

We consider next a client–server type of traffic scenario: one node ($i = 15$) behaves like a server

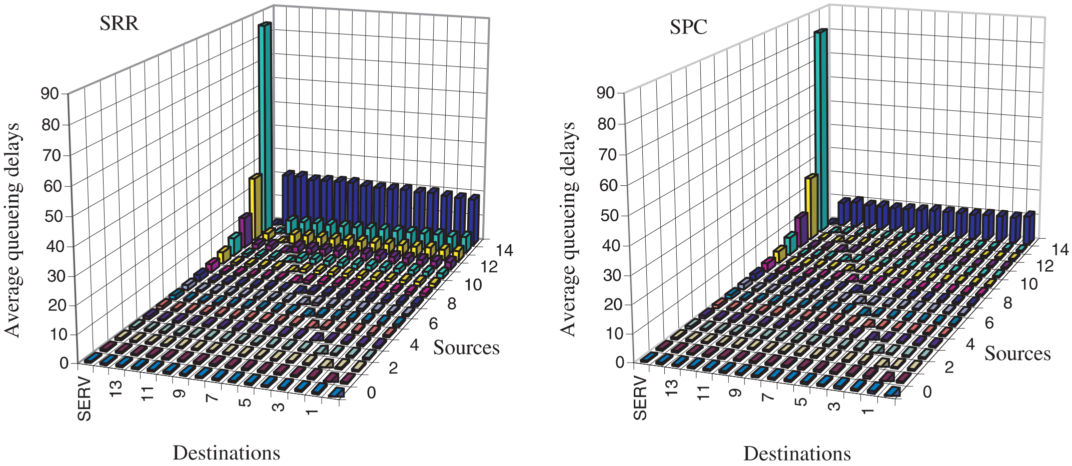

Fig. 4. Queueing delays for SRR and SPC in a client–server traffic scenario.

generating and attracting a large fraction of the network throughput; the remaining $M-1$ nodes behave like clients and direct half of their traffic to the server node, while the other half of the traffic is uniformly distributed among client nodes. The server generates an amount of traffic that is equal to half the total traffic generated by clients, and evenly distributes it to client nodes. The arrival rates are, therefore, $\lambda_{15} = \Lambda/3$ and $\lambda_i = (2/3)\Lambda/(M-1)$ $\forall i \neq 15$, while $p_{15,j} = 1/(M-1)$ $\forall j \neq 15$, $p_{ij} = 1/[2(M-2)]$ $\forall i,j \neq 15,\ i \neq j$, and $p_{i,15} = 1/2$ $\forall i \neq 15$.

An improvement in packet delay performance at the server node (see row for Source 15) is achieved with respect to SRR by using either of the two a posteriori access strategies, as it can be seen in Fig. 4, where the delay performance of SRR and SPC are reported for all source/destination pairs with a network load $\Lambda = 0.95$. A posteriori strategies allow the server to optimize its transmission scheduling, while SRR forces the server to behave like in a TDMA scheme, which is optimal only under overload. Results for LQC are very similar to SPC.

Looking now at destination 15 (labeled SERV in the figure), no significant differences are found on how the strategies treat packets transmitted by the lowest priority clients and directed to the server. The poor access performance of lowest priority clients on the channel leading to the server are mainly due to the lack of free slots on that channel (whose load is 0.9), therefore also a posteriori strategies are not able to improve that situation.

A strong reduction in the delays at client nodes with low priority on the channel leading to the server, i.e., source nodes 12, 13 and 14, when transmitting toward other client nodes is observed for the a posteriori strategy, as shown in Fig. 4. Under SRR, client nodes close to the server attempt to transmit packets to other clients only once for each SRR frame, when the corresponding transmission queues are selected according to the SRR preferential scheduling algorithm. This occurs because, every time the queue scheduled by SRR is found empty, a transmission is attempted on the server channel, where the transmission queue is very long. As a consequence, a packet destined to other clients must wait, on average, at least half the SRR frame before being successfully transmitted. Under SPC (and LQC), every time the server channel is found busy, a packet directed to other clients can be transmitted successfully. Thus, the waiting time of packets directed to other clients is shorter.

4.4. Other non-uniform traffic scenarios

Several other traffic scenarios were considered, in order to exhaustively investigate the differences between a posteriori access strategies and SRR.

We discuss some relevant results obtained under three traffic scenarios.

In the first scenario, only the eight nodes with even index are active. Each active node generates the same amount of traffic that is evenly transmitted towards the other active nodes.

In this case, the a posteriori access strategies manage to fully exploit the network capacity whereas SRR fails, as documented in Fig. 5, that plots node throughput versus node load. In each node, once every two slots the transmission queue selected according to the SRR preferential scheduling is found empty, and the transmission of the packet at the head of the longest queue is attempted. As a consequence, also in heavily loaded conditions (i.e., when queues are almost never empty) contentions among nodes are likely to arise.

In the second traffic scenario only three nodes are active, i.e., nodes 0, 1, and 2. Node 0 transmits all its packets to node 3, node 1 transmits all its packets to node 4, and finally, node 2 transmits half of its packets to node 3 and half of its packets to node 4. Nodes 0 and 1 generate the same amount of traffic $\lambda_0 = \lambda_1 = 0.5$, while $\lambda_2 = 2\lambda_0 = 1.0$.

In this case, both SRR and a posteriori strategies are not able to fully exploit the network capacity; to obtain this goal, all the slots on both channels leading to nodes 3 and 4 must be continuously filled by the three transmitting nodes. Nodes 0 and 1 should alternate their transmissions, so as to fill just one channel in each time slot, leaving the other channel empty for transmissions of node 2. In this case, node 2 would be able to transmit in each slot on the free channel.

However, either under SRR or any of the a posteriori strategies, transmissions of nodes 0 and 1 are completely independent; thus, with probability 0.25 nodes 0 and 1 transmit in the same time slot in two different channels. Node 2 finds both channel busy, and cannot transmit any packet; with probability 0.25, instead, both channels are left free from nodes 0 and 1, in the same time slot; since node 2 is able to fill only one slot at a time, being equipped with a single transmitter, this leads to an obvious throughput inefficiency. In addition, under SRR, node 2 is not always able to transmit even when a free channel is left by nodes 0 and 1, since it could make the wrong a priori choice with probability 0.5. The effect of these phenomena is shown in Fig. 6, where throughput experimented by node 2 on the channels leading to nodes 3 and 4 are reported for the three considered access strategies. SRR performs worse than both a posteriori strategies, but the maximum achievable throughput is never reached.

Finally, we considered several randomly generated traffic patterns. In Fig. 7 we report the loads for each source–destination pair for one of those patterns, the global network load being $\Lambda = 14$; in the left hand of the picture, the aggregate loads by source and by destination (i.e. the sum over the rows and columns) are also reported. It can be observed that some nodes generate, on average, more than one packet per slot. In these overload

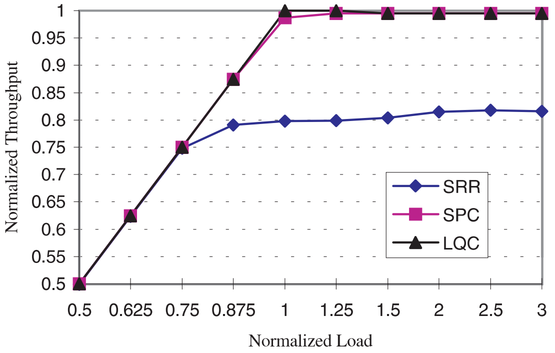

Fig. 5. Throughput for SPC, LQC and SRR when only 8 nodes are active.

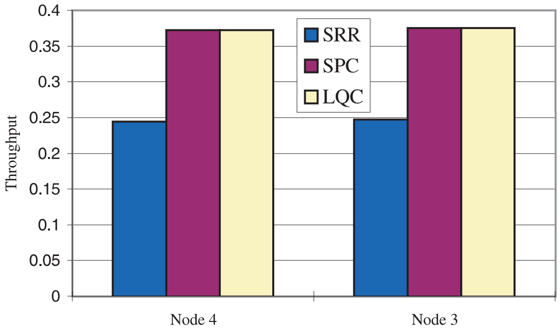

Fig. 6. Throughput for SRR, SPC and LQC at node 2 on the two channels leading to nodes 3 and 4.

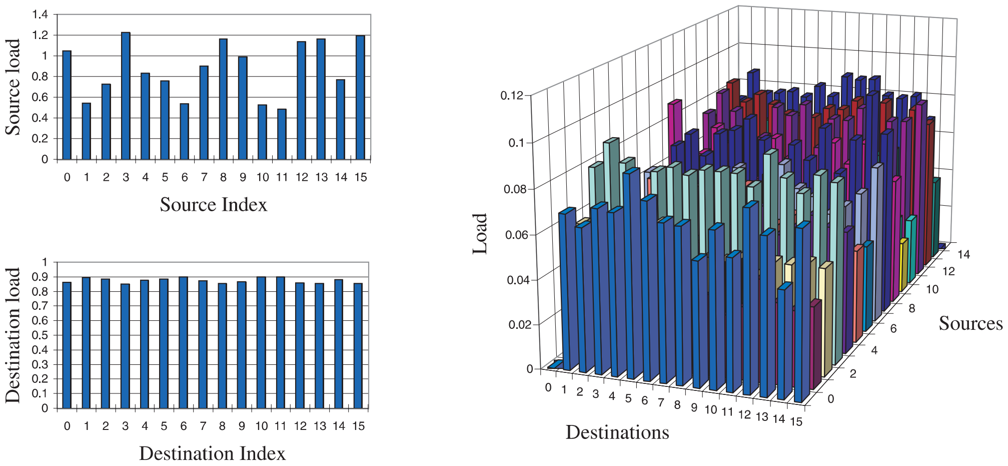

Fig. 7. Source–destination loads and average load by source and by destination.

conditions, nodes experiment losses in transmission queues, since the generated traffic exceeds their transmission capacity. Delays are a function of the transmission queue length, and are roughly equal to the product between the buffer size and the average access time, i.e., the queue service time; throughput are the reciprocal of the access time.

In Fig. 8 delays at node 3, 4, 9 and 11 are shown for the SRR, SPC and LQC strategies under traffic load reported in Fig. 7. Destinations are ordered by decreasing priority, similarly to Fig. 2. Focusing on node 3, whose load is equal to 1.223, it can be observed that the delays experienced under the SPC and LQC strategies are slightly smaller than

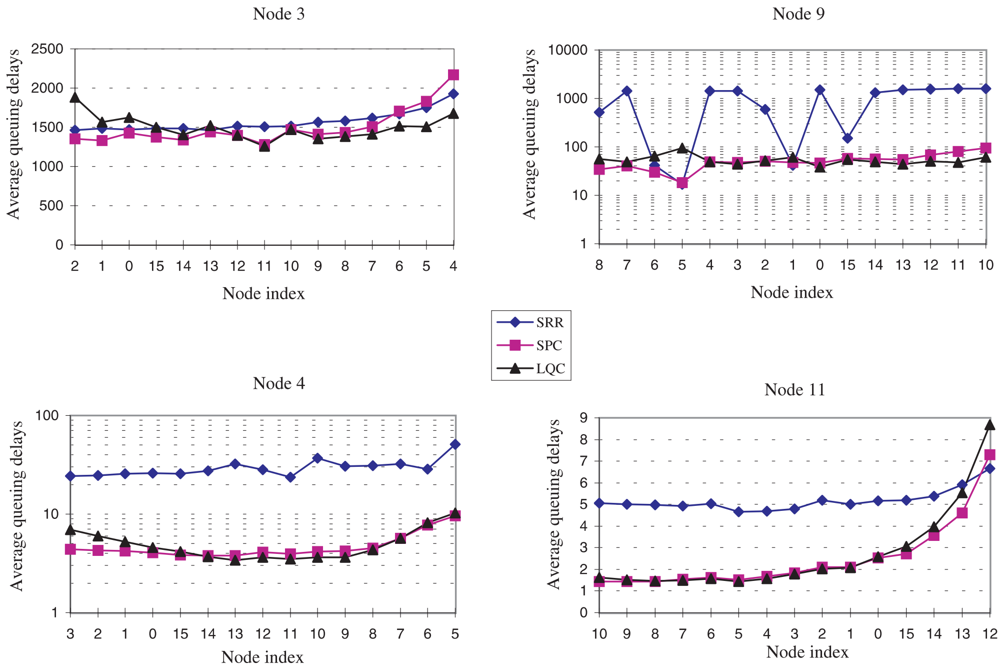

Fig. 8. Queuing delays at node 3, 4, 9 and 11, under the randomly generated traffic scenario of Fig. 7.

the ones experimented under SRR for almost all priorities; thus, throughput are slightly larger. This is again due to the fact that, with a posteriori strategies, the node is able to fully exploit the network transmission capacity; under SRR, conversely, it may contend for the channel with other nodes, thus being forced to refrain from transmission to avoid contention.

For the queueing delays at nodes 4, 9, 11, whose loads are 0.835, 0.991 and 0.485, respectively, a significant reduction is observed when SPC or LQC are used (with the only exception of the lowest priority packets of node 11). The stronger improvement is observed at node 9 (note the logarithmic vertical scale), whose load approaches 1.

All other randomly generated patterns confirm these observations.

5. Conclusions

This paper studied the performance versus complexity trade-offs in two families of access strategies used in MAC protocols for slotted multi-channel all-optical ring networks whose nodes have a distinct transmission queue for each channel and limited data transfer capabilities, i.e., one tunable transmitter and one fixed receiver.

The first family, termed a priori, comprises those strategies that select the packet for transmission based on the distribution of the packets in the transmission queues. The second family, termed a posteriori, consists of strategies that select the packet for transmission after having sensed the channels. Both families operate solely on local information, i.e., network nodes are not aware of other nodes' state. Nodes have a "carrier sense" capability to inspect the channel availability in order to avoid packet collisions in the channels.

We proposed two novel a posteriori strategies, SPC and LQC, that require an increase in the hardware and algorithmic complexity with respect to a priori strategies, but lead to significant performance improvements under several traffic patterns. Beyond these expected performance gains, the main outcome of our simulation experiments is that a posteriori strategies are more robust than a priori strategies with respect to irregular traffic patterns and ring latencies. Among the two a posteriori strategies, LQC, being based on a simpler algorithm and leading essentially to the same performance, seems the more promising. Our conclusion is that the increase of complexity required by a posteriori access strategies when compared to a priori strategies is well justified when the network under design is expected to operate under highly unbalanced traffic scenarios.

Previous works [11,26,27] showed that the SRR access strategy alone, and in general any a priori strategy, does not guarantee by itself the full exploitation of the network bandwidth when the load on some channel exceeds the channel capacity. To design a more effective MAC protocol, a fairness control scheme can be superimposed on SRR to achieve a fair partition of the network resources among nodes [11,26]. Moreover, a priori strategies, even with fairness control, lack the ability to provide acceptable service to applications requiring a certain degree of QoS; to overcome this limitation, in [27] a reservation scheme was proposed, which permits to efficiently transport real time traffic in the slotted multi-ring. The same fairness control and reservation schemes can be used in conjunction with the a posteriori access strategies presented in this paper, leading to a higher-performance, yet implementable, MAC protocol for WDM rings.

Acknowledgements

This work was supported in part by the Italian National Research Council, and by the Italian Ministry for University and Scientific Research. A preliminary version of this paper was presented at IEEE Globecom 1998 in Sydney, Australia.

References

[1] C.A. Brackett et al., A scalable multiwavelength multihop optical network: a proposal for research on all-optical networks, IEEE/OSA Journal of Lightwave Technology 11 (5/6) (1993) 736–753.

[2] B. Mukherjee, WDM-based local lightwave networks part I: single-hop systems, IEEE Network 6 (1992) 12–27.

[3] M.-S. Chen, N.R. Dono, R. Ramaswami, A media-access protocol for packet-switched wavelength division multiac-

cess metropolitan area networks, IEEE Journal on Selected Areas in Communications 8 (6) (1990) 1048–1057.

[4] I. Chlamtac, A. Fumagalli, L.G. Kazovsky, P.T. Poggiolini, A multi-Gbit/s WDM optical packet network with physical ring topology and multi-subcarrier header encoding, ECOC '93, Montreux, Switzerland, September 1993; also in: Journal of High Speed Networks 4 (2) (1995) 201–219.

[5] E. Modiano, R. Barry, Design and analysis of an asynchronous WDM local area network using a master/slave scheduler, INFOCOM '99, New York, March 1999, pp. 900–907.

[6] M. Ajmone Marsan, A. Fumagalli, E. Leonardi, F. Neri, P. Poggiolini, Daisy: a scalable all-optical packet network with multifiber ring topology, Computer Networks & ISDN Systems 30 (1998) 1065–1082.

[7] Y. Cai et al., Demonstration of photonic packet-switched ring networks with optically transparent nodes, IEEE Photonic Technology Letters 6 (9) (1994) 1139–1141.

[8] M.J. Chawki et al., Wavelength reuse scheme in a WDM unidirectional ring network using a proper fiber grating add/drop multiplexer, OFC '95, San Diego, CA, February 1995.

[9] H. Toba, K. Oda, K. Inoue, K. Nosu, T. Kitoh, An optical FDM-based self-healing ring network employing arrayed waveguide grating filters and EDFA's with level equalizers, IEEE Journal on Selected Areas in Communications 14 (5) (1996) 800–813.

[10] T.-H. Wu, Fiber Network Survivability, Artech House, Boston, 1992.

[11] M. Ajmone Marsan, A. Bianco, E. Leonardi, M. Meo, F. Neri, MAC protocols and fairness control in WDM multirings with tunable transmitters and fixed receivers, IEEE/OSA Journal of Lightwave Technology 14 (6) (1996) 1230–1244.

[12] M. Ajmone Marsan, A. Bianco, E. Leonardi, M. Meo, F. Neri, On the capacity of MAC protocols for all-optical WDM multi-rings, INFOCOM'96, San Francisco, CA, March 1996.

[13] M. Ajmone Marsan, A. Bianco, E. Leonardi, F. Neri, S. Toniolo, An almost optimal MAC protocol for all-optical WDM multi-rings with tunable transmitters and fixed receivers, ICC'97, Montreal, Canada, May 1997.

[14] F. Delorme, Widely tunable 1.55-μm lasers for wavelength-division-multiplexed optical fiber communication, IEEE Journal of Quantum Electronics 34 (9) (1998) 1706–1716.

[15] A. Budman et al., Multigigabit optical packet switch for self-routing networks with subcarrier addressing, OFC'92, paper Tu04, San José, CA, February 1992.

[16] P. Poggiolini, S. Benedetto, Theory of subcarrier encoding of packet headers in quasi-all-optical broadband WDM networks, IEEE/OSA Journal of Lightwave Technology 12 (10) (1994) 1869–1881.

[17] M. Karol, M. Hluchyj, S. Morgan, Input versus output queuing on a space division switch, IEEE Transactions on Communications 35 (12) (1987) 1347–1356.

[18] Y. Awdeh Ra'ed, H.T. Mouftah, Survey of ATM switch architectures, Computer Networks & ISDN Systems 27 (1995) 1567–1613.

[19] N. McKeown, A. Mekkittikul, A practical scheduling algorithm to achieve 100% throughput in input-queued switches, IEEE INFOCOM'98, San Francisco, CA.

[20] M. Ajmone Marsan, A. Bianco, E. Filippi, P. Giaccone, E. Leonardi, F. Neri. On the behavior of input queuing switch architectures, European Transactions on Telecommunications (ETT) 10 (2) (1999) 111–124.

[21] R.E. Tarjan, Data Structures and Network Algorithms, Society for Industrial and Applied Mathematics, Pennsylvania, November 1983.

[22] S. Even, O. Kariv, An $O(n^{2.5})$ algorithm for maximum matching in general graphs, 16th Symposium on Foundations of Computer Science, 1975, pp. 100–112.

[23] R. Schoenen, G. Post, G. Sander, Weighted arbitration algorithms with priorities for input-queued switches with 100% throughput, 3rd IEEE International Workshop on Broadband Switching Systems (BSS'99), Kingston, Ontario, Canada, June 1999.

[24] A. Bianco, A. Fumagalli, E. Leonardi, F. Neri, S. Toniolo, Multicast traffic in a WDM ring based ATM switch, EUROPTO International Symposium on Advanced Imaging and Network Technologies, Berlin, Germany, October 1996 (SPIE Proceedings vol. 2953).

[25] M. Ajmone Marsan, E. Leonardi, M. Meo, F. Neri, Modeling slotted WDM rings with discrete-time Markovian models, Computer Networks & ISDN Systems 32 (5) (2000) 599–615.

[26] M. Ajmone Marsan, A. Bianco, E. Leonardi, F. Neri, S. Toniolo, MetaRing fairness control schemes in all-optical WDM rings, INFOCOM'97, Kobe, Japan, April 1997.

[27] M. Ajmone Marsan, A. Bianco, E. Leonardi, A. Morabito, F. Neri, All-optical WDM multi-rings with differentiated QoS, IEEE Communications Magazine 2 (37) (1999) 58–66.

Andrea Bianco is an Assistant Professor at the Electronics Department of Politecnico di Torino, in Italy, since December 1994. He was born in Torino, Italy, in 1962. He holds a Dr. Ing. degree in Electronics Engineering since 1986 and a Ph.D. in electronics Engineering since 1993, both from Politecnico di Torino. Since 1994 he was an Assistant Professor at Politecnico di Torino, first in the Dipartimento di Sistemi di Produzione, later in the Dipartimento di Elettronica. In 1993 he visited Hewlett-Packard Labs in Palo Alto, CA. His current research interests are in the fields of protocols for all-optical networks and switch architectures for high-speed networks. A. Bianco is a Member of IEEE.

Andrea Fumagalli received his Ph.D. from the Politecnico di Torino, Italy, in 1992. From 1987 to 1989, he worked as a consultant for the national telephone research center in Italy, CSELT, where he was involved in the design and implementation of advanced telephone services supported over a fiber-based packet-switching ring network, LION. Between 1990 and 1992, he was a visiting researcher at the University of Massachusetts, where he defined and studied the concept of switched delay lines used as a storage method for all-optical packet-switching. From 1992 to 1998, he was on the faculty of the Electrical Engineering Department at the Politecnico di Torino. In 1994 and 1995, on leave of absence from the Politecnico di Torino, he was at the University of Massachusetts where he was responsible for a number of projects sponsored by NSF and DARPA. Since 1997, he has been Associate Professor of Electrical Engineering at the University of Texas at Dallas and member of the Center for Advanced Telecommunications Systems and Services (CATSS). Dr. Fumagalli's research involves architectures of optical and wireless networks, and performance evaluation of such networks. He has worked in optical and high speed communications since 1987, supported by several national and international funding agencies, including ESPRIT, EIT, NSF, DARPA and DoD. His publications include about seventy papers in refered journals and conferences. Dr. Fumagalli has been involved in a number of professional activities, including guest editorship for the special issue on ATM Networks of the Journal on Communications; the Journal of High Speed Networks, as a guest editor for the special issue on Optical Networks; and the European Transactions on Telecommunications, as the guest editor for the special issue on WDM Networks. He is the Associate Editor of the SPIE/Baltzer Optical Networks Magazine.

Vincenzo Distefano received a Dr. Ing. degree in Electronics Engineering in 1997 from Politecnico di Torino, with a thesis on access protocols for WDM rings.

Emilio Leonardi is an Assistant Professor at the Electronics Department of Politecnico di Torino since 1999. He was born in Cosenza, Italy, in 1967. He received a Dr. Ing. degree in Electronics Engineering in 1991 and a Ph.D. in Telecommunications Engineering in 1995, both from Politecnico di Torino. In 1995, he spent one year at the Computer Science Department of the University of California, Los Angeles (UCLA), where he was involved in the Supercomputer-SuperNet (SSN) project aimed at the design of an high capacity optical network architecture. During the summer 1999, he joined the High-Speed Research Group of Lucent Technologies, in Holmdel, NJ, where he worked on the design of scheduling algorithms for high capacity switch architectures. His research interests are in the fields of all-optical networks, switching architectures, and queueing theory. E. Leonardi is a Member of IEEE.

Fabio Neri is an Associate Professor at the Electronics Department of Politecnico di Torino, Turin, Italy. He was born in Novara, Italy, in 1958. He received his Dr. Ing. and Ph.D. degree in Electrical 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). In the summer of 1998 he visited Bell Laboratories / Lucent Technologies in Holmdel, NJ. 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. He has co-authored over 50 papers published in international journals and presented in leading international conferences. F. Neri is a member of IEEE.