

# Wavelength Reuse for Efficient Packet-Switched Transport in an AWG-Based Metro WDM Network

Michael Scheutzow, Martin Maier, Martin Reisslein, *Member, IEEE*, and Adam Wolisz, *Senior Member, IEEE*

**Abstract**—Metro wavelength-division multiplexed (WDM) networks play an important role in the emerging Internet hierarchy; they interconnect the backbone WDM networks and the local-access networks. The current circuit-switched SONET/synchronous digital hierarchy (SDH)-over-WDM-ring metro networks are expected to become a serious bottleneck—the so-called metro gap—as they are faced with an increasing amount of bursty packet data traffic and quickly increasing bandwidths in the backbone networks and access networks. Innovative metro WDM networks that are highly efficient and able to handle variable-size packets are needed to alleviate the metro gap. In this paper, we study an arrayed-waveguide grating (AWG)-based single-hop WDM metro network. We analyze the photonic switching of variable-size packets with spatial wavelength reuse. We derive computationally efficient and accurate expressions for the network throughput and delay. Our extensive numerical investigations—based on our analytical results and simulations—reveal that spatial wavelength reuse is crucial for efficient photonic packet switching. In typical scenarios, spatial wavelength reuse increases the throughput by 60% while reducing the delay by 40%. Also, the throughput of our AWG-based network with spatial wavelength reuse is roughly 70% larger than the throughput of a comparable single-hop WDM network based on a passive star coupler (PSC).

**Index Terms**—Arrayed-waveguide grating (AWG), medium access control, metro wavelength-division multiplexed (WDM) network, multiple free spectral ranges, passive star coupler (PSC), photonic packet switching, spatial wavelength reuse.

## I. INTRODUCTION

THE Internet of the future may be viewed as a three-level hierarchy consisting of backbone networks, metro networks, and access networks. Backbone networks will provide almost infinite bandwidth based on wavelength-division multiplexing (WDM) links. These WDM links are connected with reconfigurable all-optical add-drop multiplexers (OADMs) and all-optical cross connects (OXC)s controlled by multiprotocol lambda switching (MPλS) [1], optical burst switching (OBS) [2], and optical packet switching (OPS) [3]–[5] mechanisms. Access networks transport data to (and from) individual users.

Manuscript received June 13, 2002; revised March 10, 2003. This work was supported in part by the Federal German Ministry of Education and Research within the TransiNet Project, the DFG Research Center “Mathematics for key technologies” (FZT86) Berlin, and by the National Science Foundation under Grant CAREER ANI-0133252. A shorter version of this paper has appeared in *Proc. IEEE INFOCOM*, New York, June 2002, pp. 1432–1441 [47].

M. Scheutzow is with the Department of Mathematics, Technical University of Berlin, 10587 Berlin, Germany (e-mail: ms@math.tu-berlin.de).

M. Maier and A. Wolisz are with the Telecommunication Networks Group, Technical University of Berlin, 10587 Berlin, Germany (e-mail: maier@ee.tu-berlin.de; wolisz@ee.tu-berlin.de).

M. Reisslein is with the Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-7206 USA (e-mail: reisslein@asu.edu).

Digital Object Identifier 10.1109/JLT.2003.812723

By employing advanced local-area network (LAN) technologies, such as Gigabit Ethernet, broad-band access, such as xDSL and cable modems, as well as high-speed next-generation wireless systems, such as UMTS, access networks provide an ever increasing amount of bandwidth. Metro networks interconnect the high-speed WDM backbone networks and the high-speed access networks. Current metro networks are typically SONET/synchronous digital hierarchy (SDH)-over-WDM rings which are based on circuit switching and carry the ever-increasing amount of bursty data traffic only inefficiently. In addition, content providers increasingly place proxy caches in metro networks. These proxies further increase the load on metro networks. Metro networks are therefore expected to become a serious bottleneck—the so-called metro gap—in the future Internet. For these reasons, there is an urgent need for innovative metro network architectures and protocols [6].

Two key requirements for metro networks are 1) flexibility, and 2) efficiency. Flexibility is required since metro networks have to support a wide range of heterogeneous protocols, such as asynchronous transfer mode (ATM), Frame Relay, SONET/SDH, and Internet protocol (IP). This requires, in particular, that the metro networks are able to transport packets of *different sizes*. Efficiency is required because metro networks are highly cost sensitive. Therefore, the deployed WDM networking components and the WDM networking resources (in particular, wavelengths) must be utilized efficiently. As we demonstrate in this paper, a crucial technique for achieving high efficiency is *spatial wavelength reuse*. By spatial wavelength reuse we mean that in our arrayed-waveguide grating (AWG)-based metro WDM network (outlined in Section II) all wavelengths are used at all AWG ports simultaneously.

This paper builds on earlier work [7], in which we have proposed a novel AWG-based single-hop WDM network, which provides a dramatically increased degree of concurrency by i) using multiple free spectral ranges (FSRs), ii) spatially reusing all wavelengths at each AWG port, and iii) exploiting spreading techniques to enable simultaneous transmission of control and data. This earlier work focused primarily on the network architecture and the medium access control (MAC) protocol. The elementary analysis conducted in [7] provided very limited insights into the performance of the proposed network. The performance analysis in [7] is limited in that it considered only fixed-size packets and did not consider spatial wavelength reuse. However, the efficient transmission of variable-size packets is of paramount importance for future metro networks. In this paper we study 1) the photonic switching of packets of different sizes, and 2) the spatial wavelength reuse in the AWG-based network proposed in earlier work. The main contribution of this paper is to develop a stochastic model to evaluate the performance

of photonic packet switching with spatial wavelength reuse in the AWG-based network. Our analytical model gives computationally efficient and accurate expressions for the throughput and delay in the network. Our numerical results indicate that spatial wavelength reuse is crucial for efficient photonic packet switching. For typical scenarios, spatial wavelength reuse increases the throughput by 60% while reducing the delay by 40%.

This paper is organized as follows. In the following subsection, we give a quick overview of related work. In Section II, we briefly review the architecture of the studied AWG-based metro network as well as the reasoning for selecting this architecture. In Section III, we briefly review the MAC protocol for the studied network. In Section IV, we develop a stochastic model for the performance evaluation of the transmission of variable-size packets with spatial wavelength reuse. This model and performance evaluation are our main contributions. In Section V, we use our analytical results to conduct numerical investigations. We also conduct simulations to verify the accuracy of our analytical results. In Section VI, we report on additional extensive simulations that i) examine the backoff in our MAC protocol, ii) evaluate the network performance for larger scheduling windows and nodal buffers, and iii) compare the throughput performance of our AWG-based network with that of a passive star coupler (PSC)-based network. We conclude the paper in Section VII.

#### A. Related Work

Single-hop metro WDM networks based on the PSC have been studied extensively, see, for instance, [8]–[18] as well as the surveys [19], [20]. The studied networks are so-called *broadcast-and-select* networks. Each transmission is broadcast over the PSC to all nodes. The intended recipient processes the transmission, whereas the other nodes ignore the transmission. The primary limitation of the PSC-based networks is that each wavelength provides only one communication channel between any pair of network nodes at any point in time. Wavelengths, however, are scarce, especially with the low-cost coarse WDM technology that is used in cost-sensitive metro networks. By partitioning the network into several PSC-based clusters and interconnecting the clusters in some fashion, the same set of wavelengths can be reused at each cluster PSC [21]–[23].

The network studied in this paper is fundamentally different from the PSC-based networks in that it exploits the wavelength-routing property of the AWG. In the studied AWG-based network, transmissions are not broadcast. Instead, they are selectively forwarded by using the appropriate wavelength, as discussed in detail later. This results in a *switched* network as opposed to the aforementioned broadcast networks. In addition, in our network, wavelengths are spatially reused at the different AWG input ports, which allows for a high degree of concurrency and efficient use of the scarce wavelength resources.

Apart from a few older studies, such as [8], [24], [25], metro WDM networks have just recently begun to attract the interest from the research community [26]. A metro network based on OADMs has recently been studied in [27]. This network is geared toward optical *circuit* switching. In [28], a ring network

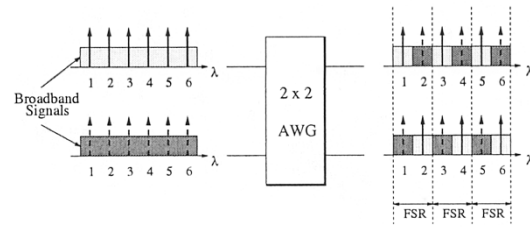


Fig. 1. Wavelength routing in AWG.

employing dynamic wavelength add-drop multiplexers is studied. Slotted packet-switched ring networks are studied in [29], [30]. In [31], a ring WDM network with AWG based OADMs has been proposed whose frequency-cyclic nature can be used for easily upgrading the network capacity. For interconnecting such WDM rings in an efficient and cost-effective way, the AWG was used as a passive wavelength-routing hub in [32]. General design principles and architectures for networks based on AWGs are studied in [33]–[38]. On-line scheduling algorithms for an AWG-based network with static wavelength assignment and static time-division multiple access, which is fundamentally different from the on-demand reservations in our network, are studied in [39]. A protection routing strategy employing AWGs is developed in [40] and a packet switch based on an AWG is studied in [41].

The HORNET metro network [42], [43] allows for optical *packet* switching. HORNET [42], [43] and the metro networks studied in [27]–[31] have a ring topology, i.e., transmissions typically have to traverse multiple network nodes. These networks are therefore fundamentally different from the packet-switched single-hop network studied here. We note that AWG-based single-hop WDM networks are also studied in [44]–[46]. However, these networks have either a more complex hub structure with wavelength converters and a centralized resource management [44], [45] or require a large number of transceivers at each node [46]. In contrast, our network is completely *passive* and resources are allocated in a *distributed* fashion (as described in the subsequent sections). Moreover, *multiple* FSRs of the underlying AWG are used to increase the degree of concurrency and thereby to improve the throughput-delay performance significantly. Each node is equipped with one *single* tunable transceiver and an off-the-shelf broad-band light source [e.g., light-emitting diode (LED)]. Finally, the novel node architecture of the proposed metro network allows for simultaneous transmission of data and control *without* requiring a separate control channel or an additional receiver at each node. Therefore, all wavelengths can be used for data transmission.

We note that a summary of the results of this paper has appeared in a shorter conference paper [47]. The conference paper states only the final analytical result of the throughput-delay analysis and provides numerical results on the impact of a subset of the system parameters. In contrast, in this extended paper, we present the derivation of the analytical results and examine the impact of all system parameters numerically. Furthermore, this extended paper provides additional simulation results on the MAC protocol backoff, the impact of scheduling window and nodal buffers, as well as a comparison with a PSC-based network.

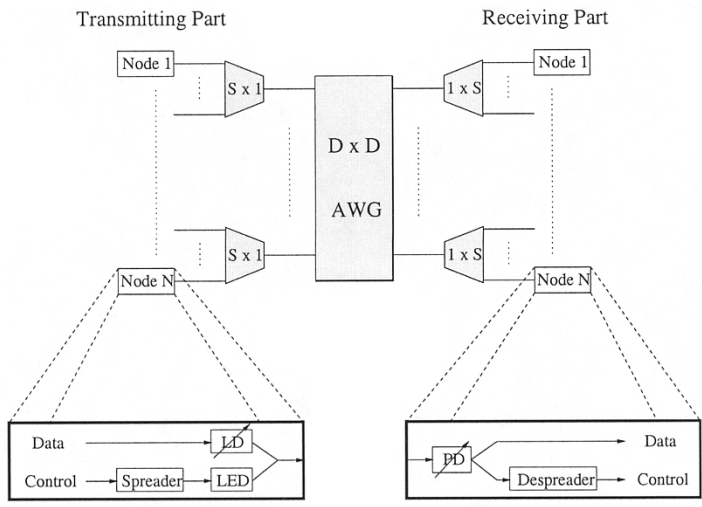


Fig. 2. Network and node architecture.

II. ARCHITECTURE

In this section, we outline the architecture of the studied AWG based network. First, we briefly review the physical properties of the AWG. The AWG [48]–[50] is a passive wavelength-routing device with a wide range of applications [51], [52]. In our network, we use the AWG as wavelength router [53]. We note that the recent development of AWGs with a crosstalk as low as  $-40$  dB makes it possible to spatially reuse all wavelengths at each AWG input port simultaneously, which, in turn, increases the capacity of AWG-based networks significantly. Moreover, by deploying *athermal* AWGs which do not require any temperature control, the costs and management of AWG-based networks can be reduced significantly [54].

The wavelength routing properties of the AWG are illustrated in Fig. 1, where six wavelengths [from laser diodes (LDs)] are sent into each input port of an AWG with degree  $D = 2$ . The AWG routes every second wavelength to the same output port. This period of the wavelength routing is referred to as *free spectral range (FSR)*. In the depicted example, we exploit  $R = 3$  FSRs, each consisting of  $D = 2$  wavelengths. Note that each FSR provides one wavelength channel for communication between a given AWG input port and a given AWG output port. In addition to the LD wavelengths, a broad-band signal (e.g., from a LED) which spans all six wavelength channels is sent into each AWG input port. As illustrated in Fig. 1, the AWG slices the broad-band signal and routes the slices to the respective output ports. Thus, spectral slicing can be used to broadcast low-speed information, e.g., control information, from a given input port to all output ports.

The considered metro WDM network is based on a  $D \times D$  AWG, as shown in Fig. 2. At each AWG input port, a wavelength-insensitive  $S \times 1$  combiner collects data from  $S$  attached nodes. Similarly, at each AWG output port signals are distributed to  $S$  nodes by a wavelength-insensitive  $1 \times S$  splitter (note that these splitters can also be used for optical multicasting). Each node is composed of a transmitting part and a receiving part. The transmitting part of a node is attached to one of the combiner ports. The receiving part of the same node is located at

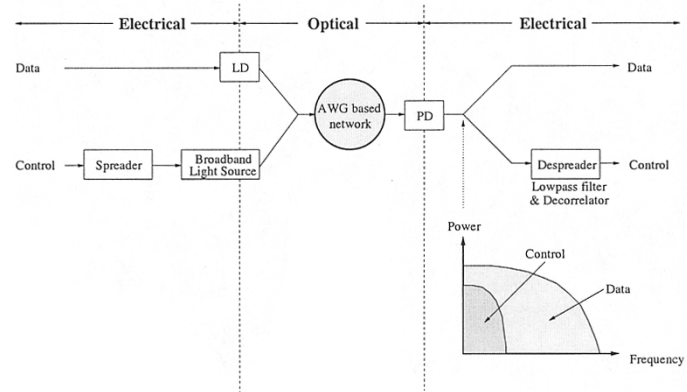


Fig. 3. Illustration of spectral spreading of control information.

the opposite splitter port. Thus, the network connects  $N = D \cdot S$  nodes. Each node contains a tunable LD and a tunable photodiode (PD) for data transmission and reception, respectively. In addition, each node uses a broad-band light source, e.g., LED, for broadcasting control packets by means of spectral slicing. The control information is spread in the electrical domain by means of direct sequence spread-spectrum techniques [55], [56]. Both data and control signals are combined and then passed through the AWG-based network, as illustrated in Fig. 3. At the output of the network, a PD is tuned to the same wavelength as the LD. The PD detects the LD wavelength and corresponding slice of the broad-band signal and converts the combined optical signal into the electrical domain. The modulation speed and launch power of the broad-band signal are such that the control signal has i) a smaller bandwidth, and ii) a smaller power level than the data signal [57], as illustrated in the lower right corner of Fig. 3. Thus, the control signal does not significantly distort the data signal. The data signal is received without requiring any further processing (except for possibly some high-pass filtering). The control signal is retrieved by low-pass filtering the combined data and (spread) control signal. The filtered signal is despread by a decorrelator which multiplies the signal with the corresponding spreading sequence. The spreading of the control

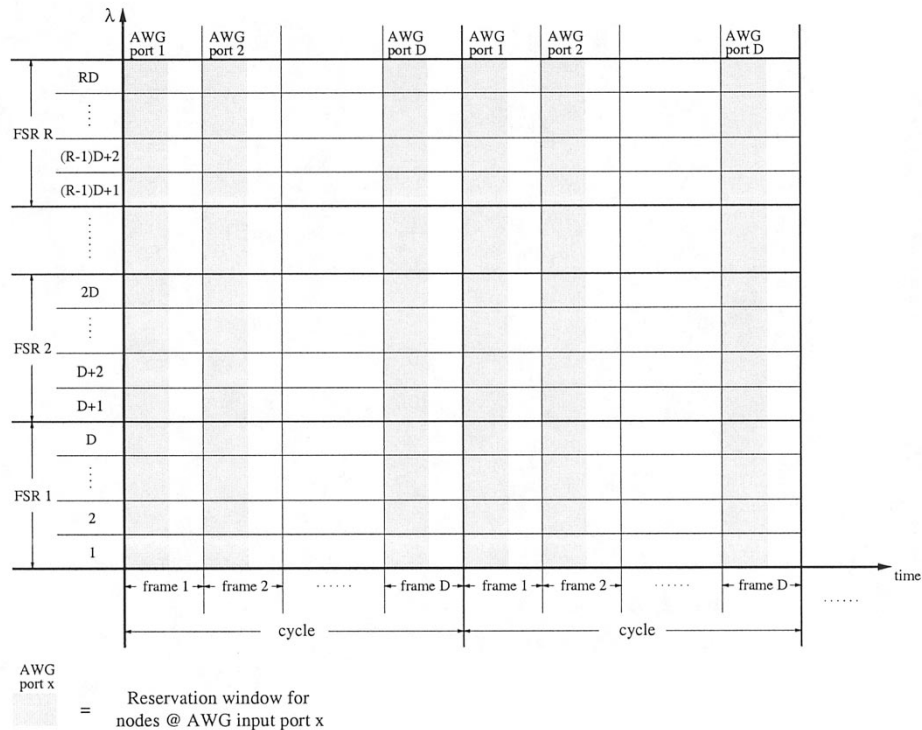


Fig. 4. Wavelength assignment and timing structure of MAC protocol at a given AWG input port.

signal has the advantage that data and control are sent simultaneously (i.e., there is no time division between signaling and data transmission), resulting in increased bandwidth efficiency.

### III. MAC PROTOCOL

In this section, we give a brief overview of our MAC protocol; we refer the interested reader to [7] for more details. In our network, each node has a tunable transmitter (TT) and a tunable receiver (TR). This TT-TR structure allows for high flexibility in the data transmissions and receptions and has the potential to achieve load balancing, as well as improved channel utilization and throughput-delay performance. However, with TT-TR nodes not only channel collisions but also receiver collisions may occur. Typically, a MAC protocol is employed to arbitrate the access to the wavelengths and thus to mitigate collisions. Generally, MAC protocols for single-hop WDM networks fall into the three main categories of i) preallocation protocols, ii) random-access protocols, and iii) reservation (pretransmission coordination) protocols, comprehensively surveyed in [20]. (Since MAC protocols for AWG-based networks have received little attention so far, the survey [20] studies the large body of literature on MAC protocols for PSC-based networks. Some key learned lessons from this literature, however, are considered generally valid and guide the design of the MAC protocol for our AWG-based network.) Preallocation protocols statically assign a wavelength to a node during a periodically recurring time slot. Preallocation generally gives high utilization only for uniform nonbursty traffic and is thus poorly suited for the bursty traffic in future metro networks. Random-access protocols do not require any preallocations to nodes or coordination among nodes. For medium to high traffic

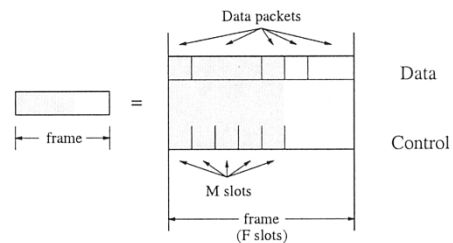


Fig. 5. Frame format.

loads, however, collisions become very frequent resulting in small throughput and large delay. Reservation protocols employ pretransmission coordination (reservation signaling) to assign wavelengths and receivers on demand. With the so-called attempt-and-defer type of reservation protocol, data packets are only transmitted after a successful reservation. Thus, attempt-and-defer protocols completely avoid channel and receiver collisions. This approach is generally preferable in a TT-TR system with bursty traffic and we adopt it for our data packet transmissions. For the pretransmission coordination we transmit small control packets according to a random-access protocol, namely, a modified slotted ALOHA protocol. This approach is adopted since i) random-access control packet transmission, as opposed to fixed assignments, makes the network scalable, and ii) for the typical large propagation delay to control packet transmission delay ratio, slotted ALOHA is superior to carrier-sensing based access.

In our MAC protocol, time is divided into *cycles*, as illustrated in Fig. 4. Each cycle consists of  $D$  frames. Each frame contains  $F$  slots, as illustrated in Fig. 5. The slot length is equal to the transmission time of a control packet. Each frame is partitioned

into the first  $M$ ,  $1 \leq M < F$ , slots and the remaining  $(F - M)$  slots. In the first  $M$  slots (the shaded area in Figs. 4 and 5), control packet transmissions take place simultaneously with data packet transmissions which do not exploit spatial wavelength reuse. In order to avoid receiver collisions of control packets, the receivers at all nodes must be tuned (locked) to one of the LED slices carrying the control information during these first  $M$  slots. Owing to the routing characteristics of the AWG, only nodes attached to the same combiner can send control packets in a given frame. Nodes attached to different AWG input ports send their control packets in different frames. Specifically, all nodes attached to AWG input port  $o$ ,  $1 \leq o \leq D$  (via a common combiner) send their control packets in frame  $o$  of the cycle. Thus, nodes can transmit control packets only in one frame per cycle. During the first  $M$  slots of frame  $o$ , control and data packets can be transmitted simultaneously by the nodes attached to AWG input port  $o$ . Due to the routing characteristics of the AWG, transmissions from the other AWG input ports cannot be received during this time interval. Consequently, the transmitters at the other ports cannot send data during this time interval. Thus, wavelengths *cannot* be spatially reused during this time interval. Therefore, data packets that are longer than  $(F - M)$  slots can be transmitted only in those frames in which the corresponding source node is allowed to send its control packets. In the last  $(F - M)$  slots of each frame no control packets are sent. The receivers are unlocked, allowing transmission between *any* pair of nodes in *each* frame provided that the corresponding data packet is not longer than  $(F - M)$  slots. This allows for spatial wavelength reuse—the main focus of this paper. Note that in general, long data packets (length  $> F - M$  slots) are harder to schedule than short ones (length  $\leq F - M$  slots).

When a data packet arrives to a node attached to AWG input port  $o$ , the node's LED broadcasts the corresponding control packet in one of the first  $M$  slots of the frame assigned to AWG input port  $o$ . The control packet has four fields (see [7]): Destination address (unicast/multicast), length, type (packet/circuit switched), and forward error correction (FEC). The control packet is transmitted on a contention basis using a modified version of slotted ALOHA. Note that using a derivative of slotted ALOHA keeps the control packet contention simple, which is of paramount importance in very-high-speed networks. Furthermore, not assigning each node a dedicated reservation slot makes the network scalable and allows for new nodes to join the network without service disruptions. Every node (including the sending node) collects all control packets by locking its receiver to one of the LED slices carrying the control information during the first  $M$  slots of every frame. Thus, each node maintains global knowledge of all the other nodes' activities (and also learns whether its own control packet collided in the control packet contention or not). Note that this approach introduces only a one-way end-to-end propagation delay (i.e., half the round-trip time). Whereas the conventional approach with control packets and explicit acknowledgment introduces the two-way end-to-end propagation delay (i.e., the full round-trip time). Also, note that the control packet broadcast completely avoids receiver collisions of control packets and allows each node to learn immediately whether the control packet suffered a channel collision. If a control packet collides,

it is retransmitted in the next cycle with probability  $p$ ; with probability  $(1 - p)$ , the retransmission is deferred by one cycle. The successfully received control packets are processed by all nodes in a distributed fashion. Each node applies the same scheduling algorithm and thus comes to the same conclusion. The scheduling algorithm tries to schedule the data packets within the scheduling window of  $D$  frames (i.e., one cycle). If the scheduling fails, the source node retransmits the control packet. Given the very-high-speed nature of optical networks and that each node has to process all other nodes' control packets, we employ a simple first-come first-served and first-fit scheduling policy to avoid a computational bottleneck.

## IV. ANALYSIS

### A. System and Traffic Model

In our analysis, we consider a system with a large  $S$ . Our analysis is approximate for finite  $S$  and exact in the asymptotic limit  $S \rightarrow \infty$ . Throughout our analysis we assume that the propagation delay is no larger than one cycle (this is reasonable for a metropolitan area network). Thus, if a control packet is sent in a given frame, the corresponding data packet could be scheduled for transmission one cycle later. We assume that all nodes are equidistant from the AWG, i.e., the propagation delay is the same for all nodes (which is easily achieved with low-loss fiber delay lines in real networks).

We assume that each node has a buffer that can hold a single data packet and a single control packet. We make the following assumptions about the traffic-generation process. Suppose that a node's control packet has just been 1) successfully transmitted, and 2) the corresponding data packet has been successfully scheduled (within the scheduling window of one cycle; see Section IV-C). With probability  $\sigma$  this node then generates the control packet for the next data packet right before the beginning of the next frame in which the node can send the next control packet (i.e., one cycle after the previous control packet was sent). If no control packet is generated, then the node waits for one cycle and then generates a new control packet with probability  $\sigma$ , and so on. The node's buffer may hold the scheduled (but not yet transmitted) data packet and the next control packet at the same time. This next control packet is sent with probability one in the next frame assigned to the node's AWG input port (possibly simultaneously with the scheduled data packet). A data packet is purged from the node's buffer at the end of the frame during which it is transmitted. After a data packet is purged from the buffer, the next data packet is placed in the buffer, provided the corresponding control packet is already in the buffer.

If a control packet fails in the slotted ALOHA contention or the data packet scheduling, then the node retransmits the control packet in the next frame assigned to the node's AWG input port with probability  $p$ , with probability  $(1 - p)$  it defers the retransmission by one cycle. In this next cycle, the node transmits the control packet with probability  $p$  and defers the transmission with probability  $(1 - p)$ , and so on. We define

$$\tilde{\alpha} := \frac{S\sigma}{M} \quad \text{and} \quad \alpha := \frac{Sp}{M}. \quad (1)$$

We conduct an approximate analysis for large  $S$ . Our analysis becomes asymptotically exact when  $S \rightarrow \infty$  and  $\tilde{\alpha}$  as well as  $\alpha$  (and also  $M$ ) are fixed [with  $\sigma$  and  $p$  chosen so as to satisfy (1)].

We consider uniform unicast traffic. A data packet is destined to any one of the  $N$  nodes (including the sending node, for simplicity) with equal probability  $1/N$ . Let  $L$  denote the length of a data packet in slots. A data packet is long (has size  $L = F$ ) with probability  $q$ , i.e.,  $P(L = F) = q$ . A data packet is short (has size  $L = K$ ,  $1 \leq K \leq F - M$ ) with probability  $(1 - q)$ , i.e.,  $P(L = K) = 1 - q$ . Our model can be extended to more complex packet size distributions at the expense of introducing more notation and a more complex arbitration policy in Section IV-C.

If a control packet fails (either in the slotted ALOHA or the scheduling), the size of the corresponding data packet is not changed. However, we do assume nonpersistence for the destination in our analysis; i.e., a new random destination is drawn for each attempt of transmitting a control packet. We note that our comparisons with simulations with persistent destinations (see Section V) clearly show that the analysis provides very accurate results despite the destination node nonpersistence assumption. We also note that our analysis can be extended to persistent destinations in a straightforward manner by maintaining the random variables introduced in the following for each AWG output port.

Now consider the nodes attached to a given (fixed) AWG input port  $o$ ,  $1 \leq o \leq D$ . These nodes send their control packets in frame  $o$  of a given cycle. We refer to the nodes that at the beginning of frame  $o$  hold an old packet, that is, a control packet that has failed in slotted ALOHA or scheduling, as “old.” We refer to all the other nodes as “new.” Note that the set of “new” nodes comprises both the nodes that have generated a new (never before transmitted) control packet as well as the nodes that have deferred the generation of a new control packet. Let  $\eta$  be a random variable denoting the number of “new” nodes at AWG input port  $o$ , and let

$$\nu := \frac{E[\eta]}{S}. \quad (2)$$

Let  $\lambda_l$  be a random variable denoting the number of nodes at port  $o$  that are to send a control packet corresponding to a long data packet next (irrespective of whether a given node is “old” or “new,” and keeping in mind that the set of “new” nodes also comprises those nodes that have deferred the generation of the next control packet; those nodes are accounted for in  $\lambda_l$  if the next generated control packet corresponds to a long data packet). Let  $\lambda_k$  be a random variable denoting the number of nodes at port  $o$  that are to send a control packet corresponding to a short data packet next. By definition,  $\lambda_k = S - \lambda_l$ . Let

$$\tilde{q} := \frac{E[\lambda_l]}{S} \quad (3)$$

denote the expected fraction of long packets to be sent. We expect that  $\tilde{q}$  is typically larger than  $q$  since long packets are harder to schedule and, thus, typically require more retransmissions (of control packets).

## B. Analysis of Control Packet Contention

First, we calculate the number of control packets from nodes attached to AWG input port  $o$ ,  $1 \leq o \leq D$ , that are successful in the slotted ALOHA contention in frame  $o$ . Let  $Y_i^n$ ,  $i = 1, \dots, M$ , be a random variable denoting the number of control packets that were randomly transmitted in slot  $i$ ,  $i = 1, \dots, M$ , by “new” nodes. Recall that each of the  $\eta$  “new” nodes sends a control packet with probability  $\sigma$  in the frame. Thus

$$P(Y_i^n = k) = \binom{\eta}{k} \left(\frac{\sigma}{M}\right)^k \left(1 - \frac{\sigma}{M}\right)^{\eta-k}, \quad k = 0, 1, \dots, \eta. \quad (4)$$

Throughout our analysis, we assume that  $S$  is large and that  $\tilde{\alpha}$  and  $\alpha$  are fixed. We may, therefore, reasonably approximate the BIN( $\eta$ ,  $\sigma/M$ ) distribution with a Poisson ( $\eta\sigma/M$ ) distribution, that is

$$P(Y_i^n = k) \approx e^{-\eta\sigma/M} \frac{(\eta\sigma/M)^k}{k!}, \quad k = 0, 1, \dots \quad (5)$$

which is exact for  $\eta \rightarrow \infty$  with  $\eta\sigma/M$  fixed. (A refined analysis that does not approximate the binomial distribution by the Poisson distribution is given in Appendix A.) We now recall the definition  $\tilde{\alpha} := \sigma S/M$ . We also approximate  $\eta/S$  by its expectation  $\nu$ ; this is reasonable since  $\eta/S$  has only small fluctuations in steady state for large  $S$ . Thus

$$P(Y_i^n = k) \approx e^{-\tilde{\alpha}\nu} \frac{(\tilde{\alpha}\nu)^k}{k!}, \quad k = 0, 1, \dots \quad (6)$$

We note that for  $S \rightarrow \infty$ , the random variables  $Y_1^n, Y_2^n, \dots, Y_M^n$  are mutually independent. This is because a given node places with the minuscule probability  $\sigma/M$  a control packet in a given slot, say slot 1. (Note in particular that the expected value of  $Y_1^n$  is small compared to the number of “new” nodes, that is,  $\tilde{\alpha}\nu \ll \eta$ ; this is because in the considered asymptotic limit  $S \rightarrow \infty$  with  $\alpha$  fixed, we have  $1 \gg \sigma/M = \tilde{\alpha}\nu/\eta$ .) Thus,  $Y_1^n$  has almost no impact on  $Y_2^n, \dots, Y_M^n$  (see Appendix B for a formal proof).

Let  $Y_i^o$ ,  $i = 1, \dots, M$ , be a random variable denoting the number of control packets in slot  $i$ ,  $i = 1, \dots, M$ , that originate from “old” nodes. Each of the  $(S - \eta)$  “old” nodes sends a control packet with probability  $p$  in the frame. Thus

$$P(Y_i^o = k) = \binom{S - \eta}{k} \left(\frac{p}{M}\right)^k \left(1 - \frac{p}{M}\right)^{S - \eta - k}, \quad k = 0, 1, \dots, S - \eta. \quad (7)$$

Approximating this binomial ( $S - \eta$ ,  $p/M$ ) distribution by the Poisson ( $(S - \eta)p/M$ ) distribution we have

$$P(Y_i^o = k) \approx e^{-(S - \eta)p/M} \frac{[(S - \eta)p/M]^k}{k!}, \quad k = 0, 1, \dots \quad (8)$$

With  $\alpha := pS/M$  and approximating  $(S - \eta)/S$  by its expectation  $(1 - \nu)$  we get

$$P(Y_i^o = k) \approx e^{-\alpha(1 - \nu)} \frac{[\alpha(1 - \nu)]^k}{k!}, \quad k = 0, 1, \dots \quad (9)$$

We note again that for  $S \rightarrow \infty$ , the random variables  $Y_1^o, Y_2^o, \dots, Y_M^o$  are mutually independent. They are also independent of  $Y_1^n, Y_2^n, \dots, Y_M^n$ . Hence, we obtain for  $Y_i = Y_i^n + Y_i^o$

$$P(Y_i = k) \approx e^{-[\tilde{\alpha}\nu + \alpha(1-\nu)]} \frac{[\tilde{\alpha}\nu + \alpha(1-\nu)]^k}{k!}, \quad k = 0, 1, \dots \quad (10)$$

Henceforth, we let for notational convenience

$$\beta := \tilde{\alpha}\nu + \alpha(1-\nu) \quad (11)$$

i.e.,

$$P(Y_i = k) \approx e^{-\beta} \frac{\beta^k}{k!}, \quad k = 0, 1, \dots \quad (12)$$

Let  $X_i, i = 1, \dots, M$ , be a random variable indicating whether or not slot  $i$  contains a successful control packet. Specifically, let

$$X_i = \begin{cases} 1, & \text{if } Y_i = 1 \\ 0, & \text{otherwise.} \end{cases} \quad (13)$$

From (12) clearly,

$$P(X_i = 1) = \beta e^{-\beta} \quad \text{and} \quad P(X_i = 0) = 1 - \beta e^{-\beta}$$

for  $i = 1, \dots, M$ . The total number of successful control packets in the considered frame is  $\sum_{i=1}^M X_i$ , which has a binomial  $(M, \beta e^{-\beta})$  distribution, that is

$$P\left(\sum_{i=1}^M X_i = l\right) = \binom{M}{l} (\beta e^{-\beta})^l (1 - \beta e^{-\beta})^{M-l}, \quad l = 0, 1, \dots, M. \quad (14)$$

Recall from Section IV-A that each packet is destined to any one of the  $D$  AWG output ports with equal probability  $1/D$ . Let  $Z$  denote the number of successful control packets—in the considered frame—that are destined to a given (fixed) AWG output port  $d, d = 1, \dots, D$ . Clearly, from (14)

$$P(Z = k) = \binom{M}{k} \left(\frac{\beta e^{-\beta}}{D}\right)^k \left(1 - \frac{\beta e^{-\beta}}{D}\right)^{M-k}, \quad k = 0, 1, \dots, M. \quad (15)$$

(A refined approximation of  $P(Z = k)$  which does not approximate the binomial distributions of the  $Y_i^n$ 's and  $Y_i^o$ 's by Poisson distributions is given in Appendix A.) Let  $Z_l$  denote the number of successful control packets that correspond to long data packets destined to a given AWG output port  $d$ . Recall that  $\tilde{q}$  is the expected fraction of long packets to be sent. Hence

$$Z_l \sim \text{BIN}(M, \beta e^{-\beta}(1/D)\tilde{q}).$$

Similarly, let  $Z_k$  denote the number of control packets that are successful in the slotted ALOHA contention and correspond to short data packets destined to a given AWG output port  $d$ . Clearly

$$Z_k \sim \text{BIN}(M, \beta e^{-\beta}(1/D)(1 - \tilde{q})).$$

### C. Analysis of Packet Scheduling

In this subsection, we calculate the expected number of packets that are successfully scheduled. Recall from the previous subsection that the total number of long packets that 1) originate from a given AWG input port  $o, 1 \leq o \leq D$ , 2)

are successful in the slotted ALOHA contention of frame  $o$  (of a given cycle), and 3) are destined to a given AWG output port  $d, 1 \leq d \leq D$ , is  $Z_l \sim \text{BIN}(M, \beta e^{-\beta}(1/D)\tilde{q})$ . For short packets, we have  $Z_k \sim \text{BIN}(M, \beta e^{-\beta}(1/D)(1 - \tilde{q}))$ . Note that these two random variables are not independent. Let  $\mathcal{L}(S)$  be a random variable denoting the number of long (short) packets that 1) originate from a given AWG input port  $o, 1 \leq o \leq D$ , 2) are successful in the slotted ALOHA contention of frame  $o$  (of a given cycle), 3) are destined to a given AWG output port  $d, 1 \leq d \leq D$ , and 4) are successfully scheduled within the scheduling window of  $D$  frames (i.e., one cycle).

Consider the scheduling of packets from a given (fixed) AWG input port  $o$  to a given (fixed) AWG output port  $d$  over the scheduling window (i.e.,  $D$  frames). Clearly, we can schedule at most  $R$  long packets (i.e.,  $\mathcal{L} \leq R$ ) because the receivers at output port  $d$  must tune to the appropriate spectral slices during the first  $M$  slots of every frame. Thus, they can tune to a node at AWG input port  $o$  for  $F$  consecutive slots, only in the frame, during which the nodes at AWG input port  $o$  send their control packets.

Now, suppose that  $\mathcal{L} (\leq R)$  long packets are scheduled (how  $\mathcal{L}$  is determined is discussed shortly). With  $\mathcal{L}$  long packets already scheduled, we can schedule at most

$$S \leq (D - 1) \cdot R \cdot \left\lfloor \frac{F - M}{K} \right\rfloor + (R - \mathcal{L}) \left\lfloor \frac{F}{K} \right\rfloor \quad (16)$$

short packets. To see this, note that in the frame during which the nodes at AWG input port  $o$  send their control packets, there are  $(R - \mathcal{L})$  FSRs—channels between AWG input port  $o$  and AWG output port  $d$ —free for a duration of  $F$  consecutive slots. Furthermore, there are  $(D - 1)$  frames in the scheduling window during which the nodes at AWG output port  $d$  must tune (are locked) to the nodes sending control packets from the other AWG input ports for the first  $M$  slots of the frame. During each of these frames, the receivers are unlocked for  $(F - M)$  slots. The  $R$  utilized FSRs provide  $R$  parallel channels between AWG input port  $o$  and AWG output port  $d$ . Note that the  $(D - 1)R \lfloor (F - M)/K \rfloor$  component in (16) is due to the spatial reuse of wavelengths at the considered AWG input port. Without spatial wavelength reuse this component would be zero and we could schedule at most  $(R - \mathcal{L}) \lfloor F/K \rfloor$  short packets. Continuing our analysis for a network with spatial wavelength reuse, we have

$$S = \min \left\{ Z_k, (D - 1) \cdot R \cdot \left\lfloor \frac{F - M}{K} \right\rfloor + (R - \mathcal{L}) \left\lfloor \frac{F}{K} \right\rfloor \right\}. \quad (17)$$

In (17), we neglect receiver collisions, that is, we do not account for situations where a packet cannot be scheduled because its receiver is already scheduled to receive a different packet. This assumption is reasonable as receiver collisions are rather unlikely for moderately large  $S$ . This is verified by our simulations which take receiver collisions into consideration, see Section V.

In this paper, we consider a first-come first-served first-fit scheduling policy. Data packets are scheduled for the first possible slot(s) at the lowest available wavelength. To arbitrate the access to the frame which allows transmission for  $F$  contiguous slots and the  $(D - 1)$  frames which allow transmission for  $(F - M)$  contiguous slots we adopt the following *arbitration policy*. Our arbitration policy proceeds in one round if there are

$R$  or fewer successful control packets in the slotted ALOHA contention. In case there are more than  $R$  successful control packets in the slotted ALOHA contention, our arbitration policy proceeds in two rounds. First, consider the case where  $R$  or fewer control packets are successful in the slotted ALOHA contention and we have one round of arbitration. In this case, all the successful packets are scheduled in the frame with  $F$  available transmission slots. Next, consider the case where more than  $R$  control packets are successful in the slotted ALOHA contention and we have two rounds of arbitration. In this case, we scan the  $M$  slotted ALOHA slots from index 1 through  $M$ . In the first round, we schedule the first  $R$  successful packets out of the slotted ALOHA contention in the  $R$  long ( $F$  slots) transmission slots. In this round, we schedule only one packet for each of the long transmission slots, irrespective of whether the packet is long or short. At this point (having filled each of the long transmission slots with one data packet) all the remaining successful control packets that correspond to long data packets fail in the scheduling and the transmitting node has to retransmit the control packet. We then proceed with the second round. In the second round, we schedule the remaining successful control packets that correspond to short data packets. Provided  $F/K > 2$ , we schedule these short data packets for the long transmission slots that hold only one short data packet from the first round. We also schedule these short data packets for the short ( $F - M$  slots) transmission slots. After all the long and short transmission slots have been filled, the remaining short data packets fail in the scheduling and the transmitting node has to retransmit the control packet. We note that our adopted arbitration policy is just one out of many possible arbitration policies, all of which can be analyzed in a similar fashion. The first-come first-served and first-fit scheduling algorithm was chosen to meet the stringent timing requirements of reservation-based very-high-speed networks.

With the adopted arbitration policy the expected number of scheduled long packets is

$$E[\mathcal{L}] = \sum_{k=0}^M E[\mathcal{L}|Z = k] \cdot P(Z = k) \quad (18)$$

$$= \sum_{k=0}^M \min(k, R) \cdot \tilde{q} \cdot \binom{M}{k} \left(\frac{\beta e^{-\beta}}{D}\right)^k \left(1 - \frac{\beta e^{-\beta}}{D}\right)^{M-k} \quad (19)$$

$$= \tilde{q} \left\{ R - \sum_{k=0}^{\min(R, M)} \binom{M}{k} \left(\frac{\beta e^{-\beta}}{D}\right)^k \left(1 - \frac{\beta e^{-\beta}}{D}\right)^{M-k} \cdot (R - k) \right\}. \quad (20)$$

To see this, note that in case there are  $k \leq R$  successful control packets in the slotted ALOHA contention, then, on average,  $\tilde{q}k$  of these correspond to long data packets. In case  $k \geq R$ , then there are on average  $\tilde{q}R$  long packets among the first  $R$  control packets. [If the arbitration policy does not schedule the long ( $F$  slots) transmission slots first, but, say after  $l$  frames that allow transmission for  $(F - M)$  slots have been scheduled, then an expected number of  $\min\{R, \tilde{q}(lR[(F - M)/K] + R)\}$  long

data packets are scheduled given  $(\lfloor (F - M)/K \rfloor + 1)R$  or more successful control packets in the slotted ALOHA contention.] For notational convenience let

$$\varphi(\beta) := R - \sum_{k=0}^{\min(R, M)} \binom{M}{k} \left(\frac{\beta e^{-\beta}}{D}\right)^k \left(1 - \frac{\beta e^{-\beta}}{D}\right)^{M-k} \cdot (R - k). \quad (21)$$

Thus

$$E[\mathcal{L}] = \tilde{q} \cdot \varphi(\beta). \quad (22)$$

We now calculate the expected number of scheduled short packets. Generally

$$E[\mathcal{S}] = \sum_{k=0}^M E[\mathcal{S}|Z = k] \cdot P(Z = k). \quad (23)$$

First, consider the case that there are no more than  $R$  successful control packets in the slotted ALOHA contention, i.e.,  $k \leq R$ . In this case, we have with (17)

$$E[\mathcal{S}|Z = k] = E[Z_k|Z = k] \quad (24)$$

$$= (1 - \tilde{q}) \cdot k \quad (25)$$

since all the successful control packets are scheduled in the long transmission slots.

Next, consider the case  $R \leq k \leq M$ . Let  $\Theta$  denote the number of control packets that correspond to short data packets to be scheduled in the second round of arbitration. Note that  $Z_k = R - \mathcal{L} + \Theta$ , because  $(R - \mathcal{L})$  short data packets have been scheduled in the first round of arbitration. With (17) we obtain

$$\begin{aligned} E[\mathcal{S}|Z = k] &= E \left[ \min \left( R - \mathcal{L} + \Theta, (D - 1)R \left\lfloor \frac{F - M}{K} \right\rfloor \right. \right. \\ &\quad \left. \left. + (R - \mathcal{L}) \left\lfloor \frac{F}{K} \right\rfloor \right) \middle| Z = k \right] \quad (26) \end{aligned}$$

$$\begin{aligned} &= E[R - \mathcal{L}|Z = k] + E \left[ \min \left( \Theta, (D - 1)R \left\lfloor \frac{F - M}{K} \right\rfloor \right. \right. \\ &\quad \left. \left. + (R - \mathcal{L}) \left( \left\lfloor \frac{F}{K} \right\rfloor - 1 \right) \right) \middle| Z = k \right]. \quad (27) \end{aligned}$$

Note that for a network without spatial wavelength reuse, the  $(D - 1)R[(F - M)/K]$  term has to be replaced by zero in (26) and (27), as well as in all the following expressions in this subsection.

Clearly,  $E[R - \mathcal{L}|Z = k] = (1 - \tilde{q})R$ , since  $k \geq R$ . Moreover, note that conditional on  $Z = k$ ,  $k \geq R$ , the random variables  $\mathcal{L}$  and  $\Theta$  are independent. This is because the first  $R$  successful control packets in the slotted ALOHA slots determine  $\mathcal{L}$ ;  $\Theta$  is determined by the subsequent  $(k - R)$  successful control packets. Hence

$$\begin{aligned} E \left[ \min \left( \Theta, (D - 1)R \left\lfloor \frac{F - M}{K} \right\rfloor \right. \right. \\ \left. \left. + (R - \mathcal{L}) \left( \left\lfloor \frac{F}{K} \right\rfloor - 1 \right) \right) \middle| Z = k \right] \quad (28) \end{aligned}$$



$$= \sum_{j=1}^{k-R} P \left( \min \left( \Theta, (D-1)R \left\lfloor \frac{F-M}{K} \right\rfloor + (R-\mathcal{L}) \left( \left\lfloor \frac{F}{K} \right\rfloor - 1 \right) \right) \geq j \mid Z = k \right) \quad (29)$$

$$= \sum_{j=1}^{k-R} P(\Theta \geq j \mid Z = k) \cdot P \left( (D-1)R \left\lfloor \frac{F-M}{K} \right\rfloor + (R-\mathcal{L}) \left( \left\lfloor \frac{F}{K} \right\rfloor - 1 \right) \geq j \mid Z = k \right). \quad (30)$$

Now

$$P(\Theta \geq j \mid Z = k) = \sum_{m=j}^{k-R} P(\Theta = m \mid Z = k) \quad (31)$$

$$= \sum_{m=j}^{k-R} \binom{k-R}{m} (1-\tilde{q})^m \tilde{q}^{k-R-m}. \quad (32)$$

For notational convenience let

$$\gamma_j := P \left( (D-1)R \left\lfloor \frac{F-M}{K} \right\rfloor + (R-\mathcal{L}) \left( \left\lfloor \frac{F}{K} \right\rfloor - 1 \right) \geq j \mid Z = k \right) \quad (33)$$

and

$$v_j := \min \left( R, \frac{(D-1)R \left\lfloor \frac{F-M}{K} \right\rfloor - j}{\left\lfloor \frac{F}{K} \right\rfloor - 1} + R \right). \quad (34)$$

If  $\lfloor F/K \rfloor - 1 > 0$ , then

$$\gamma_j = P(\mathcal{L} \leq v_j \mid Z = k) \quad (35)$$

$$= \sum_{\{m: m \leq v_j\}} P(\mathcal{L} = m \mid Z = k) \quad (36)$$

$$= \sum_{\{m: m \leq v_j\}} \binom{R}{m} \tilde{q}^m (1-\tilde{q})^{R-m}. \quad (37)$$

In case  $\lfloor F/K \rfloor = 1$ , we have

$$\gamma_j = \begin{cases} 1, & \text{if } j \leq (D-1)R \left\lfloor \frac{F-M}{K} \right\rfloor \\ 0, & \text{otherwise.} \end{cases} \quad (38)$$

Combining the cases  $k \leq R$  and  $R \leq k \leq M$ , we obtain for  $M \geq R$ , which is typical for practical networks

$$E[\mathcal{S}] = \sum_{k=0}^R (1-\tilde{q})k \cdot P(Z=k) + (1-\tilde{q})R \cdot \sum_{k=R+1}^M P(Z=k) \quad (39)$$

$$+ \sum_{k=R+1}^M \left( \sum_{j=1}^{k-R} \gamma_j \sum_{m=j}^{k-R} \binom{k-R}{m} (1-\tilde{q})^m \tilde{q}^{k-R-m} \right) \cdot P(Z=k)$$

$$= (1-\tilde{q}) \left[ R - \sum_{k=0}^R (R-k) \cdot P(Z=k) \right] \quad (40)$$

$$+ \sum_{j=1}^{M-R} \gamma_j \sum_{m=j}^{M-R} \sum_{k=m+R}^M \binom{k-R}{m} (1-\tilde{q})^m \tilde{q}^{k-R-m} \cdot P(Z=k) =: h(\tilde{q}, \beta). \quad (41)$$

#### D. Network/System Analysis

In this subsection, we combine the analyses for the individual components of the considered network, namely, traffic model, slotted ALOHA contention, and packet scheduling. We establish two equilibrium conditions and solve for the two unknowns  $\tilde{q}$  and  $\beta$ . [Alternatively, we may consider the two unknowns  $\tilde{q}$  and  $\nu$ , noting that  $\nu = (\beta - \alpha)/(\tilde{\alpha} - \alpha)$  for  $\tilde{\alpha} \neq \alpha$ ; the case  $\tilde{\alpha} = \alpha$  is discussed at the end of this section.]

In steady state, the system satisfies the equilibrium condition

$$E[\mathcal{L}] = q(E[\mathcal{L}] + E[\mathcal{S}]). \quad (42)$$

To see this, note that in equilibrium the mean number of scheduled long packets from a given (fixed) AWG input port destined to a given (fixed) AWG output port [the left-hand side (LHS)] is equal to the mean number of newly generated long packets [the right-hand side (RHS)]. Inserting (22) and (41) in (42) gives

$$\tilde{q} \cdot \varphi(\beta) = q[\tilde{q} \cdot \varphi(\beta) + h(\tilde{q}, \beta)] \quad (43)$$

$$\Leftrightarrow (1-q) \cdot \tilde{q} \cdot \varphi(\beta) = q \cdot h(\tilde{q}, \beta). \quad (44)$$

The second equilibrium condition is

$$\frac{\sigma}{D} \cdot E[\eta] = E[\mathcal{L} + \mathcal{S}]. \quad (45)$$

This is because  $\sigma \cdot \eta$  new packets are generated in each frame at the nodes attached to a given AWG input port. With probability  $1/D$ , each of the generated packets is destined to a given (fixed) AWG output port. On the other hand,  $E[\mathcal{L} + \mathcal{S}]$  packets are scheduled (and transmitted) on average from a given AWG input port to a given AWG output port in one cycle; in equilibrium as many new packets must be generated. Inserting (1) and (2) in the LHS of (45) and (22) and (41) in the RHS of (45) we obtain

$$\frac{\tilde{\alpha} \cdot M}{D} \cdot \frac{\beta - \alpha}{\tilde{\alpha} - \alpha} = \tilde{q} \cdot \varphi(\beta) + h(\tilde{q}, \beta). \quad (46)$$

Inserting (46) into (44) we obtain

$$\tilde{q} = \frac{q \cdot \tilde{\alpha} \cdot M}{D \cdot \varphi(\beta)} \cdot \frac{\beta - \alpha}{\tilde{\alpha} - \alpha}. \quad (47)$$

Inserting (47) into (46) we obtain

$$(1-q) \cdot \frac{\tilde{\alpha} \cdot M}{D} \cdot \frac{\beta - \alpha}{\tilde{\alpha} - \alpha} = h \left( \frac{q \cdot \tilde{\alpha} \cdot M}{D \cdot \varphi(\beta)} \cdot \frac{\beta - \alpha}{\tilde{\alpha} - \alpha}, \beta \right). \quad (48)$$

We solve (48) numerically to obtain  $\beta$  [noting that by (11),  $\min(\tilde{\alpha}, \alpha) \leq \beta \leq \max(\tilde{\alpha}, \alpha)$ ]. We then insert  $\beta$  into (47) to obtain  $\tilde{q}$ . With  $\beta$  and  $\tilde{q}$ , we calculate  $E[\mathcal{L}]$  (22) and  $E[\mathcal{S}]$  (41). We define the mean throughput as the average number of transmitting nodes in steady state (which may also be interpreted as the average number of successfully transmitted data packets per frame). The mean throughput from a given (fixed) AWG input port to a given (fixed) AWG output port, i.e., the average number of nodes transmitting from a given AWG input port to a given AWG output port in steady state, is then given by

$$TH_{\text{port}} = \frac{F \cdot E[\mathcal{L}] + K \cdot E[\mathcal{S}]}{F \cdot D}. \quad (49)$$

The mean aggregate throughput of the network is

$$TH_{\text{net}} = D^2 \cdot TH_{\text{port}}. \quad (50)$$

Note that  $\mathcal{L} \leq R$  and that  $\mathcal{S}$  is bounded by the expression in (16). Thus, if  $(F - M)/K$  and  $F/K$  are integers, the aggregate throughput is bounded by  $D^2 \cdot R \cdot [1 - M/F \cdot (1 - 1/D)]$ .

We note that in case  $\tilde{\alpha} = \alpha$ , i.e.,  $\sigma = p$ , we have, from (11),  $\beta = \alpha$ . Inserting this into (44) gives an equation for  $\tilde{q}$ , which we solve numerically.

We now espouse the mean packet delay in the network. We define the mean delay as the average time period in cycles from the generation of the control packet corresponding to a data packet until the transmission of the data packet. Recall from Section IV-C that  $E[\mathcal{L}] + E[\mathcal{S}]$  is the expected number of data packets that the nodes at a given AWG input port transmit to the nodes at a given AWG output port per cycle. Now, consider a given (fixed) node  $m$ ,  $1 \leq m \leq N$ . In the assumed uniform packet traffic scenario, this node  $m$  transmits on average  $(E[\mathcal{L}] + E[\mathcal{S}])/S$  data packets to the nodes at a given AWG output port per cycle. Thus, node  $m$  transmits on average  $(E[\mathcal{L}] + E[\mathcal{S}])D/S$  data packets to the  $N$  nodes attached to the  $D$  AWG output ports per cycle. The average time period in cycles from the generation of a control packet at node  $m$  until the generation of the next control packet is, therefore,  $S/[D \cdot (E[\mathcal{L}] + E[\mathcal{S}])]$ . Note that the time period from the successful scheduling of a data packet until the generation of the control packet for the next data packet is geometrically distributed with mean  $(1 - \sigma)/\sigma$  cycles. Hence, the average delay in the network in cycles is

$$\text{Delay} = \frac{S}{D \cdot (E[\mathcal{L}] + E[\mathcal{S}])} - \frac{1 - \sigma}{\sigma} \quad (51)$$

where  $E[\mathcal{L}]$  and  $E[\mathcal{S}]$  are known from the evaluation of the throughput (49).

## V. NUMERICAL RESULTS

In this section, we show the benefit of spatial wavelength reuse and the impact of the system parameters on the throughput-delay performance of the network. Data packets can have one of two lengths. A data packet is  $F$  slots long with probability  $q$  and  $K = F - M$  slots long with probability  $(1 - q)$ . We consider the system parameters number of used FSRs  $R$ , the fraction of long data packets  $q$ , the number of reservation slots per frame  $M$ , the physical degree of the AWG  $D$ , the number of nodes  $N$ , and retransmission probability  $p$ . By default, the parameters take on the following values:  $R = 2$ ,  $q = 0.25$ ,  $M = 30$ ,  $D = 4$ ,  $N = 200$ ,  $p = 0.8$ ,  $F = 200$ , and  $K = 170$ . (For these default parameters the aggregate throughput is bounded by 28.4.) Each cycle is assumed to have a constant length of  $D \cdot F = 800$  slots. All numerical results in this section are obtained using the expression (15), which approximates the number of successful control packets in the slotted ALOHA contention by a Poisson distribution. [For a numerical evaluation of the refined approximation, that does not use the Poisson distribution, but uses directly the binomial distribution, we refer the interested reader to Appendix A. In summary, we find that using the approximate expression (15) gives very accurate results for a wide range of parameter values, as is also demonstrated by the numerical results in this section.] We also provide extensive simulation results of a more realistic network in order to verify the accuracy of the

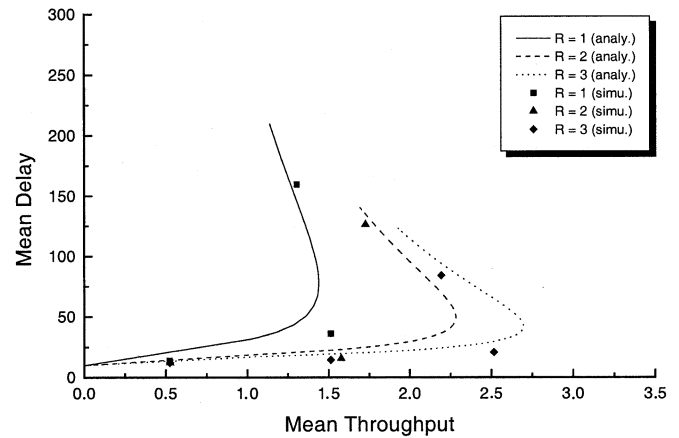


Fig. 6. Mean delay (cycles) versus mean aggregate throughput (packets per frame) for different number of used FSRs  $R \in \{1, 2, 3\}$ .

analysis. As opposed to the analysis, in the simulation a given node cannot transmit data packets to itself and both length and destination of a given data packet are not renewed when retransmitting the corresponding control packet. In addition, the simulation takes receiver collisions into account, i.e., a given data packet is not scheduled if the receiver of the intended destination node is busy. Each simulation was run for  $10^7$  slots including a warmup phase of  $10^6$  slots. Using the method of batch means we calculated the 98% confidence intervals for the mean aggregate throughput and the mean delay whose widths were less than 1% of the corresponding sample means for all simulation results.

Fig. 6 shows the mean delay versus the mean aggregate throughput as the mean arrival rate  $\sigma$  is varied from 0 to 1. As one would expect, using two FSRs instead of one (leaving all other parameters unchanged) dramatically increases the mean aggregate throughput while decreasing the mean delay. This is due to the fact that an additional FSR increases the degree of concurrency and thereby mitigates the scheduling bottleneck resulting in more successfully transmitted data packets and fewer retransmissions. Note that the number of used FSRs is limited and is determined by the transceiver tuning range, the degree  $D$  of the underlying AWG, and the channel spacing. To avoid tuning penalties we deploy fast tunable transceivers whose tuning range is typically 10–15 nm. All results presented in this section assume a channel spacing of 200 GHz, i.e., 1.6 nm at  $1.55 \mu\text{m}$ . Thus, we can use seven to ten wavelengths at each AWG input port depending on the transceiver tuning range. For all subsequent results, the number of wavelengths is set to eight. Consequently, with a  $4 \times 4$  AWG we deploy two FSRs for concurrent transmission/reception of data packets.

Figs. 7 and 8 illustrate that spatial wavelength reuse dramatically improves the throughput-delay performance of the network for variable-size data packets. Fig. 7 shows the mean aggregate throughput versus the mean arrival rate  $\sigma$  with and without spatial wavelength reuse for different fraction of long data packets  $q \in \{0, 0.5, 1.0\}$ . Simulation and analysis results match very well. For  $q = 1.0$ , i.e., all data packets have a length of  $F$  slots, the mean aggregate throughput is the same no matter whether wavelengths are spatially reused or not. This is because the data packets are too long to be scheduled in the  $(D - 1)$

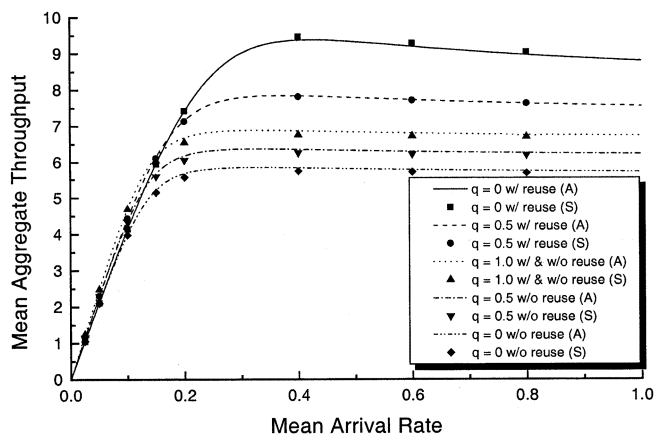


Fig. 7. Mean aggregate throughput (packets per frame) versus mean arrival rate  $\sigma$  with and without wavelength reuse for different fraction of long data packets  $q \in \{0, 0.5, 1.0\}$ .

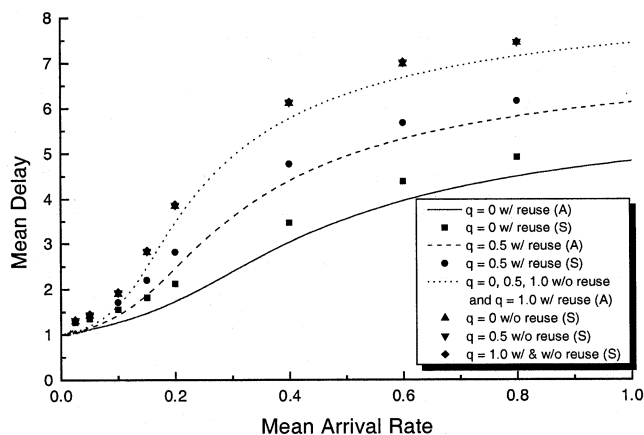


Fig. 8. Mean delay (cycles) versus mean arrival rate  $\sigma$  with and without wavelength reuse for different fraction of long data packets  $q \in \{0, 0.5, 1.0\}$ .

frames in which the corresponding nodes do not send control packets and spatial wavelength reuse would be possible in the last  $(F - M)$  slots of the frame. Thus, these frames remain unused for  $q = 1.0$ . For  $q = 0.5$ , 50% of the data packets are long ( $F$  slots) and the other 50% are short ( $K$  slots). Allowing for spatial wavelength reuse, the latter ones can now be scheduled in all frames including the aforementioned  $(D - 1)$  frames. Consequently, with wavelength reuse, more data packets are successfully transmitted, resulting in a higher throughput. In contrast, without wavelength reuse, data packets can be scheduled only in one frame per cycle in which the corresponding nodes also transmit their control packets. Furthermore, since for  $q = 0.5$  some successfully transmitted data packets are short ( $K$  slots), wavelengths are not fully utilized resulting in a lower throughput compared to  $q = 1.0$ . For  $q = 0$ , the benefit of spatial wavelength reuse becomes even more dramatic. In this case, there are only short data packets ( $K$  slots) which fill up a large number of frames leading to a further increased mean aggregate throughput. Note that for  $q = 0$ , spatial wavelength reuse significantly increases the maximum aggregate throughput by more than 60%. All curves in Fig. 7 run into saturation since for increasing  $\sigma$  no additional data packets can be scheduled due to busy channels and receivers and an increasing number of colliding control packets.

Fig. 8 depicts the mean delay versus  $\sigma$  with and without wavelength reuse for different fraction of long data packets  $q \in \{0, 0.5, 1.0\}$ . We observe that the simulation gives slightly larger delays than the analysis. This is because the simulation takes also the transmission time of data packets into account as opposed to the analysis. In the analysis, the mean delay is equal to the time interval between the generation of a given data packet and the end of the cycle in which the given data packet is successfully scheduled but not yet transmitted. All curves have in common that at very light traffic the mean delay is equal to one cycle owing to the propagation delay of the control packet. With increasing  $\sigma$  the mean delay increases due to more unsuccessful control packets. These control packets have to be retransmitted, resulting in an increased mean delay. Note that we obtain the largest delay if the aforementioned  $(D - 1)$  frames per cycle cannot be used for data transmission. This holds true not only for the cases where wavelength reuse is not deployed but also for  $q = 1.0$  with spatial wavelength reuse. This is due to the fact that for  $q = 1.0$  the data packets are too long and do not fit in the last  $(F - M)$  slots of the aforementioned  $(D - 1)$  frames. As a consequence, for these cases, fewer data packets can be successfully scheduled and the corresponding control packets have to be retransmitted more often, leading to a higher mean delay. With decreasing  $q$  there are more short data packets which can easily be scheduled in the aforementioned  $(D - 1)$  frames. Due to the resulting wavelength reuse more data packets can be successfully scheduled. Therefore, fewer control packets have to be retransmitted leading to a decreased mean delay. In particular, for  $q = 0$ , wavelengths are used very efficiently resulting in the lowest mean delay.

The impact of the number of reservation slots  $M$  per frame on the network throughput-delay performance is shown in Figs. 9 and 10. The mean aggregate throughput and the mean delay are depicted as a function of  $\sigma$  for  $M \in \{15, 20, 30, 40\}$ . Recall that by default the frame length  $F$  is set to 200 slots. Each frame is composed of  $M$  reservation slots and  $K = F - M$  slots which can be used for transmitting short packets by means of spatial wavelength reuse. Clearly, for a fixed  $F$ , increasing  $M$  decreases the length of short packets  $K$ , and reduces the contribution of the short packets to the throughput. At the same time, increasing  $M$  increases the probability of successful control packet contention. We observe from Figs. 9 and 10 that the effect of increasing the probability of successful control packet contention dominates in the considered scenario, that is, the mean throughput increases and the mean delay decreases as  $M$  increases. This indicates that the random-access reservation scheme can be a bottleneck. Indeed, for a very small number of  $M = 15$  control slots, we observe the typical bistable behavior of slotted ALOHA—the underlying mechanism of our control packet contention—in the throughput results. Note that the network throughput-delay performance could be easily improved by replacing the random access of the reservation slots with a dedicated assignment of the reservation slots. However, such a dedicated slot assignment requires reconfiguration when adding new nodes.

A given number of nodes can be connected by AWGs with different physical degree  $D$ . Figs. 11 and 12 depict for  $D \in \{2, 4, 8\}$  the mean aggregate throughput and the mean delay

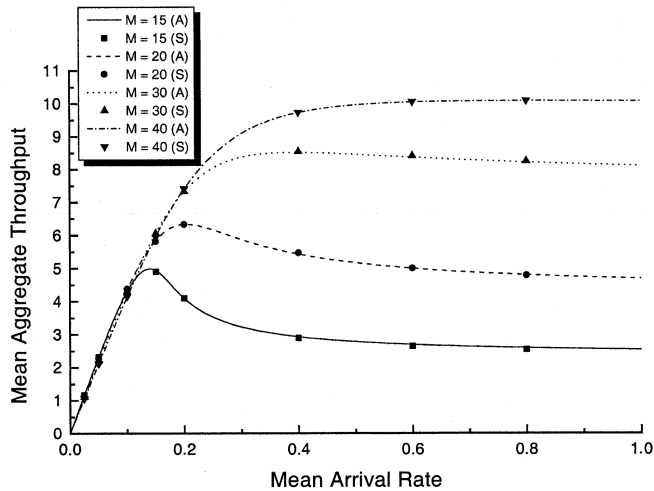


Fig. 9. Mean aggregate throughput (packets per frame) versus mean arrival rate  $\sigma$  for different number of reservation slots  $M \in \{15, 20, 30, 40\}$ .

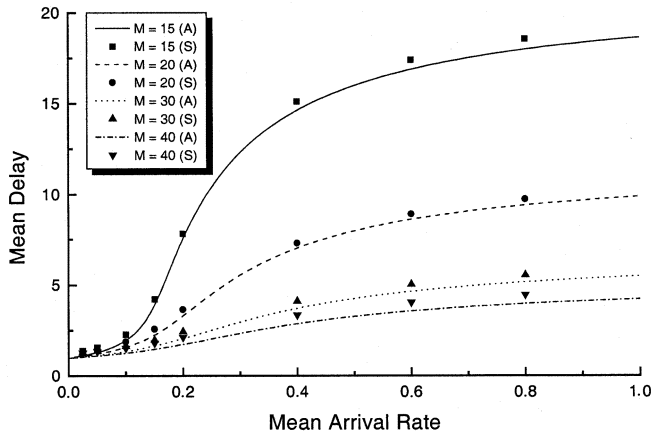


Fig. 10. Mean delay (cycles) versus mean arrival rate  $\sigma$  for different number of reservation slots  $M \in \{15, 20, 30, 40\}$ .

as a function of  $\sigma$ , respectively. Recall that we have chosen the transceiver tuning range and the channel spacing such that we make use of eight wavelengths. The number of used FSRs  $R$  is then determined only by the physical degree  $D$  of the underlying AWG and is given by  $R = 8/D$ . Consequently, for a smaller  $D$ , more FSRs are exploited, and *vice versa* for a larger  $D$ . Furthermore, for a smaller  $D$  each cycle contains fewer but longer frames, and *vice versa* for a larger  $D$ .

As shown in Fig. 11,  $D = 2$  provides the largest maximum mean aggregate throughput at light traffic. However, with increasing  $\sigma$  the mean aggregate throughput decreases. This is due to the fact that for  $D = 2$ , short data packets are rather long ( $K = 800/D - M = 370$  slots) resulting in a higher channel utilization and thereby a higher throughput at small traffic loads. But a small  $D$  also implies that for a given population  $N$ , more nodes are attached to the same combiner since  $S = N/D$ . All these  $S$  nodes make their reservations in the same frame. For an increasing  $\sigma$ , this leads to more collisions of control packets resulting in a lower mean aggregate throughput and a higher mean delay due to more retransmissions of the corresponding control packets (Fig. 12).

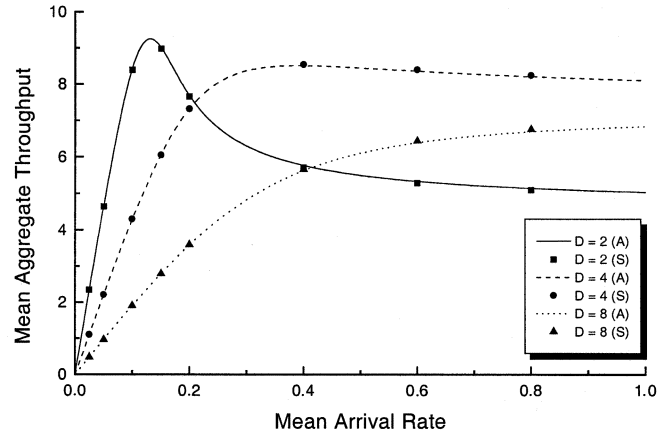


Fig. 11. Mean aggregate throughput (packets per frame) versus mean arrival rate  $\sigma$  for different AWG degree  $D \in \{2, 4, 8\}$ .

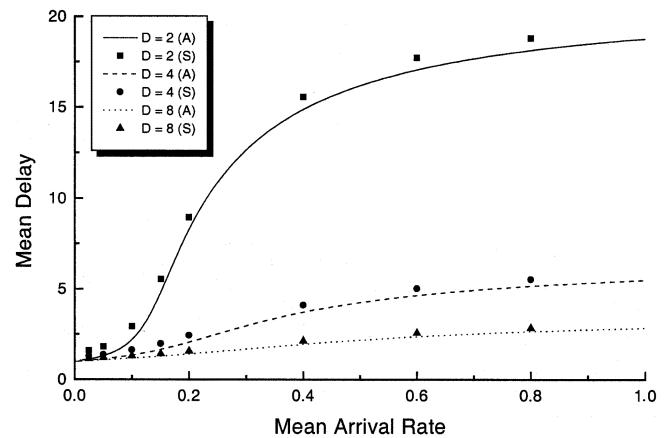


Fig. 12. Mean delay (cycles) versus mean arrival rate  $\sigma$  for different AWG degree  $D \in \{2, 4, 8\}$ .

This problem is alleviated by deploying a  $4 \times 4$  or  $8 \times 8$  AWG. For a larger  $D$ , fewer nodes send control packets in the same frame causing fewer collisions at high traffic loads. However, for  $D = 4$  and  $D = 8$  only two FSRs and one FSR can be deployed, respectively. Moreover, a larger  $D$  reduces the length of short data packets. Fig. 11 shows that for  $D = 4$  the mean aggregate throughput is rather high for a wide range of  $\sigma$ . Whereas for  $D = 8$  the throughput is rather low due to the small number of control packets per frame and the low channel utilization owing to the reduced length of short data packets. Note that for  $D = 4$  the mean aggregate throughput gradually decreases for increasing  $\sigma$ . This is because at high traffic loads control packets suffer from collisions and have to be retransmitted, resulting in a slightly higher mean delay compared to  $D = 8$ . Concluding, in terms of throughput-delay performance choosing  $D = 4$  seems to provide the best solution for a wide range of traffic loads.

Figs. 13 and 14 depict the throughput-delay performance of the network for different population  $N \in \{40, 100, 200, 300\}$ . As shown in Fig. 13, increasing  $N$  improves the mean aggregate throughput due to more reservation requests and successfully scheduled data packets. However, for  $N = 200$  and especially  $N = 300$  the throughput decreases for increasing  $\sigma$ . This

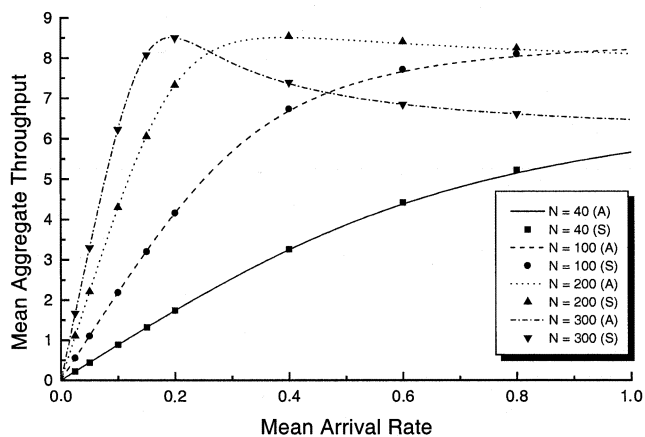


Fig. 13. Mean aggregate throughput (packets per frame) versus mean arrival rate  $\sigma$  for different population  $N \in \{40, 100, 200, 300\}$ .

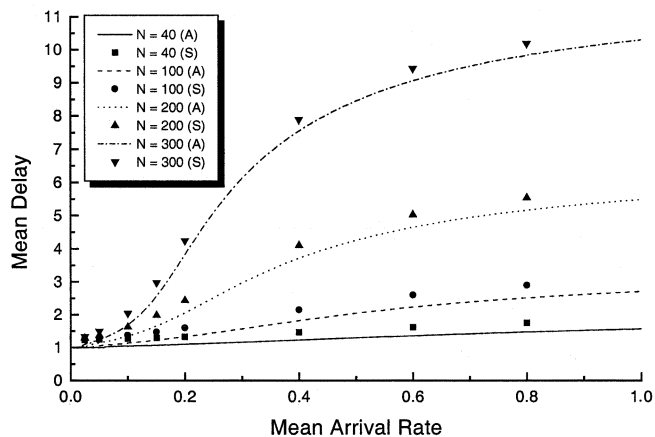


Fig. 14. Mean delay (cycles) versus mean arrival rate  $\sigma$  for different population  $N \in \{40, 100, 200, 300\}$ .

is because for large populations, more control packets suffer from channel collisions resulting in a lower mean aggregate throughput. Accordingly, this leads to higher mean delays as shown in Fig. 14. Note that simulation and analysis results match very well even for small populations despite the fact that i) we have conducted an asymptotic analysis for large  $S$ , and ii) the analysis does not take receiver collisions into account (while the simulation does). Also, the analysis assumes non-persistent destinations, whereas the destinations are persistent in the simulation.

The impact of different AWG degree  $D \in \{2, 4, 8\}$  on the system throughput-delay performance for different packet size distributions is shown in Figs. 15 and 16, respectively. Note that throughput and delay are not given as a function of  $\sigma$  but as a function of the fraction of long data packets  $q \in [0, 1]$ . Recall that we have assumed a constant cycle length of 800 slots and a fixed number of reservation slots  $M = 30$ . As a consequence, the frame length is given by  $F = 800/D$  slots and the length of short data packets is equal to  $K = F - M = 800/D - 30$  slots. Moreover, the number of used FSRs of the underlying AWG is given by  $R = 8/D$ .

Fig. 15 depicts the mean aggregate throughput versus  $q$ . For  $D = 4$ , we observe that the throughput monotonously

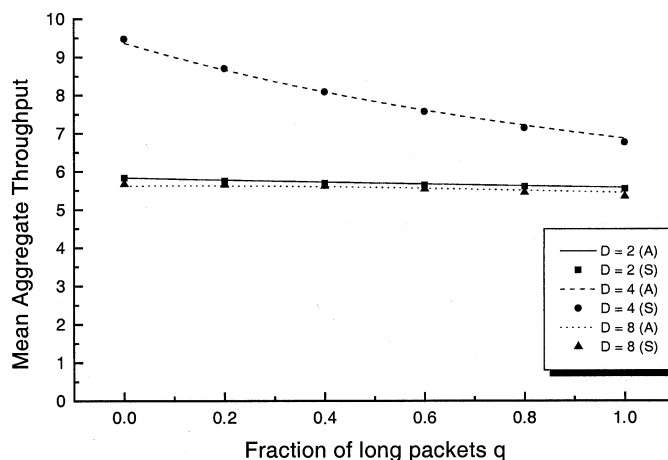


Fig. 15. Mean aggregate throughput (packets per frame) versus fraction of long data packets  $q$  for different AWG degree  $D \in \{2, 4, 8\}$ .

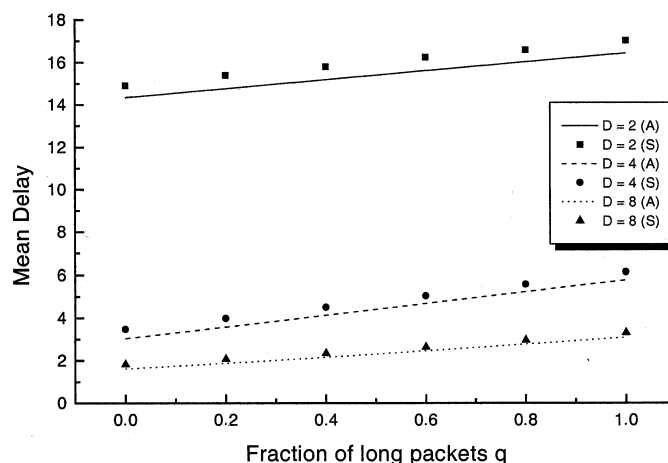


Fig. 16. Mean delay (cycles) versus fraction of long data packets  $q$  for different AWG degree  $D \in \{2, 4, 8\}$ .

decreases for increasing  $q$ . For  $q = 0$ , all data packets are short and can be scheduled in any frame resulting in a high mean aggregate throughput. For increasing  $q$ , more and more data packets are long ( $q = 1$  implies that there are only long data packets). However, long data packets can be scheduled only in one frame per cycle. In addition, at most two of them can be scheduled since  $R = 2$ . Consequently, for increasing  $q$ , fewer data packets can be scheduled resulting in a decreasing mean aggregate throughput (Fig. 15) and a higher mean delay as shown in Fig. 16. For  $D \in \{2, 8\}$ , the mean aggregate throughput is smaller than for  $D = 4$  and, more interestingly, almost independent from  $q$ . For  $D = 2$ , twice as many nodes are attached to each AWG input port compared to  $D = 4$ . As a consequence, more control packets suffer from collisions and fewer data packets are available for scheduling, resulting in a smaller mean aggregate throughput. Moreover, there are not enough control packets to fully capitalize on spatial wavelength reuse. Thus, for  $q = 0$ , only slightly more data packets are successfully scheduled than for  $q = 1$ . However, since for  $q = 1$  all successfully scheduled data packets are long as opposed to  $q = 0$ , the mean aggregate throughput is about the same in both cases. Similarly, since for  $D = 8$

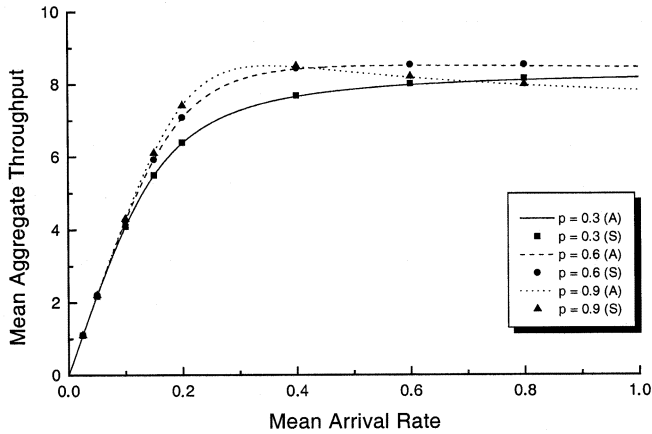


Fig. 17. Mean aggregate throughput (packets per frame) versus mean arrival rate  $\sigma$  for different retransmission probability  $p \in \{0.3, 0.6, 0.9\}$ .

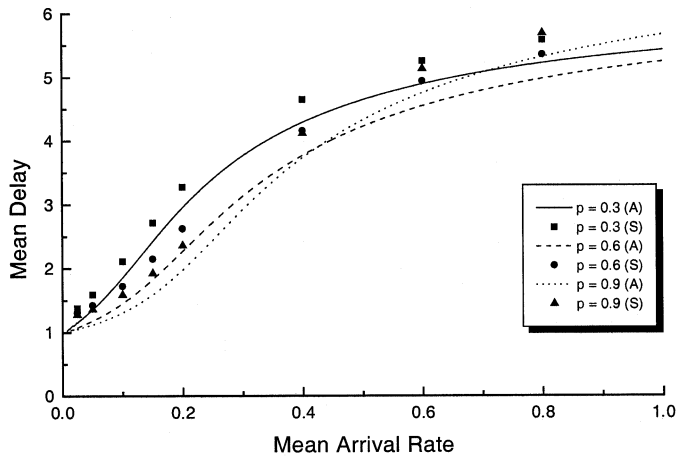


Fig. 18. Mean delay (cycles) versus mean arrival rate  $\sigma$  for different retransmission probability  $p \in \{0.3, 0.6, 0.9\}$ .

fewer nodes are attached to each AWG input port there are fewer reservation requests per frame than for  $D = 4$  resulting in a smaller throughput. However, these reservation requests experience fewer collisions significantly decreasing the mean delay, as illustrated in Fig. 16. In contrast,  $D = 2$  provides the highest mean delay due to the large number of collided control packets and their retransmissions. Note that for all  $D \in \{2, 4, 8\}$ , the mean delay grows with increasing  $q$  since for larger  $q$  fewer data packets can be scheduled owing to the lack of spatial wavelength reuse. This leads to more retransmissions of control packets and thereby to an increased mean delay. Concluding, while  $D \in \{2, 8\}$  suffers from a relatively small throughput and  $D = 2$  exhibits a too large mean delay, choosing  $D = 4$  seems to provide the best compromise in terms of throughput-delay performance.

Figs. 17 and 18 depict the throughput-delay performance of the network as a function of  $\sigma$  for different retransmission probability  $p \in \{0.3, 0.6, 0.9\}$ . As shown in Fig. 17, for  $p = 0.3$  the mean aggregate throughput grows monotonously for increasing  $\sigma$ . We observe that at high traffic loads, the slope decreases due to increasingly busy channels and transceivers. For  $p = 0.6$ , the mean aggregate throughput is larger than for  $p = 0.3$ . This is because with a larger  $p$ , nodes retransmit

collided control packets with a higher probability resulting in more successful control packets and an increased mean aggregate throughput. However, further increasing  $p$  has a detrimental impact on the throughput. For  $p = 0.9$ , nodes retransmit collided control packets after a small backoff period. As a consequence, at medium-to-high traffic loads, an increasing number of control packets collide leading to a decreased mean aggregate throughput. Fig. 18 shows that for all  $\sigma \in (0, 1]$ ,  $p = 0.3$  yields larger mean delays than  $p = 0.6$ . With a smaller  $p$ , nodes defer retransmissions of collided control packets for a larger time interval which, in turn, increases the mean delay. Note that for  $p = 0.9$ , nodes experience the smallest mean delay at light-to-medium traffic loads. At high loads, the mean delay becomes largest due to the increasing number of retransmissions of control packets.

## VI. SUPPLEMENTARY SIMULATION RESULTS

In this section, we give an overview of extensive supplementary simulations of our AWG-based network. In these simulations, we examine important additional aspects and performance metrics of the AWG-based network. We also relax some of the simplifying assumptions made in the previous analysis in order to make our investigations more realistic. Specifically, we examine the backoff of our MAC protocol, extend the scheduling window size, and equip each node with a finite buffer. In doing so, we relax the assumptions of a one-cycle scheduling window and a single-packet buffer made in our analytical model. We also conduct a benchmark comparison with a previously reported reservation protocol designed for a PSC-based single-hop metro WDM network. Due to space constraints, we give here only an overview of these investigations and refer to [58] for more details.

In the simulations, the network parameters take on the following default values:  $N = 200$ ,  $D = 2$ ,  $R = 4$ ,  $F = 200$ ,  $M = 30$ ,  $p = 1.0$ , and  $q = 0.25$ . By default, the size of the scheduling window is one cycle and each node's single-packet buffer is able to store either a long or a short data packet. Each cycle has a constant length of  $D \cdot F = 400$  slots. The simulations are conducted as described in Section V, see also [58] for more details. Throughout this section we report the 95% confidence intervals for the performance metrics.

### A. Media Access Control (MAC) Protocol Backoff

Without backoff, the retransmission probability  $p$  remains constant irrespective of how many times a given control packet has already been retransmitted. As the mean arrival rate increases, more nodes are backlogged and try to make a reservation in their assigned frame. The network becomes increasingly congested and more nodes have to retransmit their unsuccessful control packets (with constant probability  $p = 1$  in our case). This leads to an increased number of control packet collisions and retransmissions. As a result, for an increasing mean arrival rate, the mean aggregate throughput decreases while the mean delay increases dramatically, as depicted in Figs. 19 and 20, respectively.

By deploying backoff, this performance degradation is alleviated. With backoff, the retransmission probability  $p$  is reduced

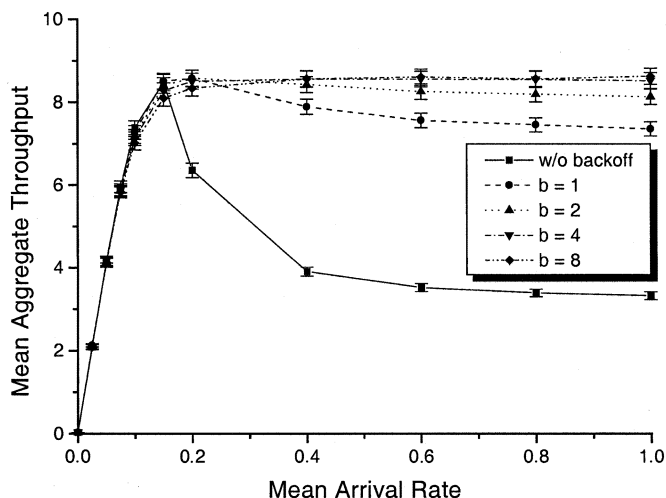


Fig. 19. Mean aggregate throughput (packets per frame) versus mean arrival rate for different backoff limits  $b \in \{1, 2, 4, 8\}$ .

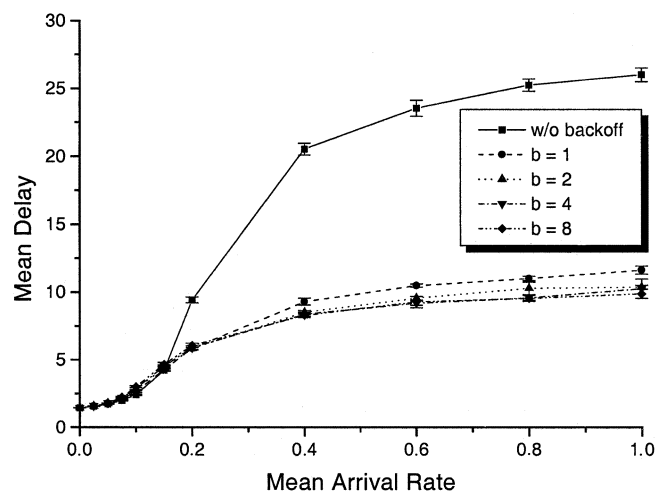


Fig. 20. Mean delay (cycles) versus mean arrival rate for different backoff limits  $b \in \{1, 2, 4, 8\}$ .

each time the reservation fails. More precisely, a given control packet which for the first time fails in making a successful reservation is retransmitted with the original retransmission probability  $p = 1$  in the next cycle. If the reservation fails again, then  $p$  is reduced by 50%. Thus, the corresponding control packet is retransmitted in the next cycle with probability  $p = 0.5$ . With probability  $(1 - p) = 0.5$ , the reservation is deferred by one cycle; in that next cycle, the control packet is retransmitted with probability  $p = 0.5$  and deferred with probability  $(1 - p) = 0.5$ , and so on. Each time the reservation fails,  $p$  is cut in half. In general,  $p$  is reduced by 50% at most  $b$  times, where  $b \geq 1$  denotes the backoff limit. Once  $p$  has been halved  $b$  times, the control packet is retransmitted with probability  $p_{\text{orig}}/2^b$  (where  $p_{\text{orig}}$  denotes the original retransmission probability of one in our case) in each of the following cycles until the reservation is successful, i.e., there is no limit on the number of attempts.

Figs. 19 and 20 show the positive impact of backoff on the throughput-delay performance of the network for  $b \in \{1, 2, 4, 8\}$ . We observe that already with  $b = 1$ , i.e.,  $p$

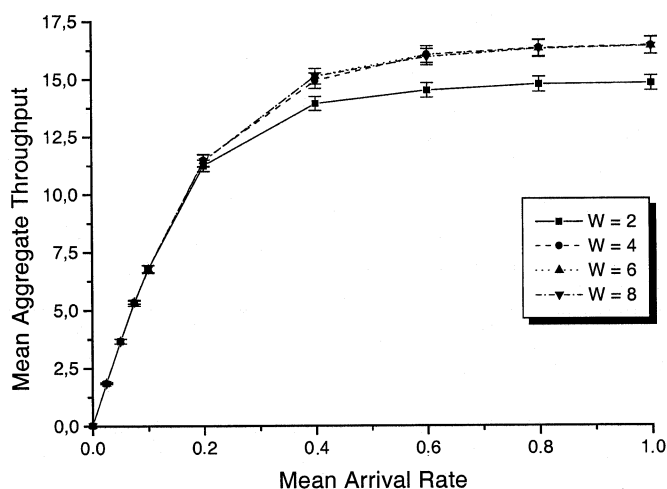


Fig. 21. Mean aggregate throughput (packets per frame) versus mean arrival rate for  $b = 4$ ,  $M = 60$ ,  $R = 8$ , and different scheduling window lengths  $W \in \{2, 4, 6, 8\}$  frames.

is halved only once, the mean aggregate throughput is significantly increased for medium-to-high traffic loads, while the mean delay is considerably decreased. We observe from Fig. 19 that with  $b = 4$ , the mean aggregate throughput does not decrease for an increasing mean arrival rate. This property of MAC protocols is known as *self-stability*. Further increasing  $b$  does not yield a better throughput-delay performance. Therefore, we set  $b = 4$  in the subsequent simulations.

In further simulations that we cannot include here due to space constraints, we have examined the setting of the number of reservation slots per frame  $M$  and the number of used FSRs  $R$  for the MAC protocol with backoff, see [58] for details. In brief, we find that the mean delay is reduced for larger  $M$ . The throughput, on the other hand, first increases for increasing  $M$ , reaches a maximum around  $M = 60$ , and then drops off as  $M$  is increased further. We, therefore, set  $M = 60$  in the subsequent simulations. We also found that increasing the number of used FSRs  $R$  increases the throughput-delay performance significantly. However, for a given transceiver tuning range  $\Lambda = D \cdot R$  and AWG degree  $D$ , a larger  $R$  results in a smaller channel spacing. Generally, AWGs with a smaller channel spacing exhibit a larger crosstalk. In order to achieve acceptable crosstalk values we set  $R = 8$  in the subsequent simulations. For  $D = 2$  and a typical fast transceiver tuning range of 12 nm, setting  $R = 8$  translates into a channel spacing of 100 GHz.

### B. Scheduling Window and Nodal Buffer

Fig. 21 depicts the mean aggregate throughput as a function of the mean arrival rate for different scheduling window size  $W \in \{2, 4, 6, 8\}$ , where  $W$  is given in frames. We observe that the throughput is increased when the scheduling window is enlarged from two frames (i.e., one cycle for  $D = 2$ ) to four frames (i.e., two cycles for  $D = 2$ ). Further increasing  $W$  has no significant impact, as the network resources are almost fully utilized for  $W \geq 4$ . Similar observations hold for the mean delay which is decreased for larger  $W$ , see [58]. For the subsequent simulations we set  $W = 8$ . This is because in the following we will also consider the case  $D = 4$ ; in order to benefit from a scheduling

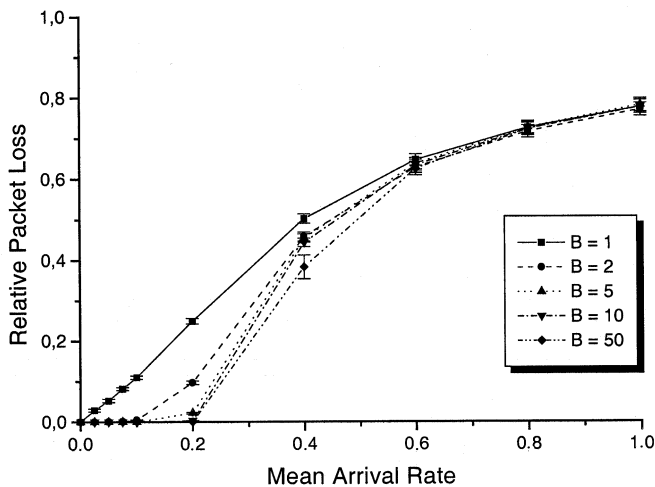


Fig. 22. Relative packet loss versus mean arrival rate for  $b = 4$ ,  $M = 60$ ,  $R = 8$ ,  $W = 8$ , and different buffer sizes  $B \in \{1, 2, 5, 10, 50\}$  packets.

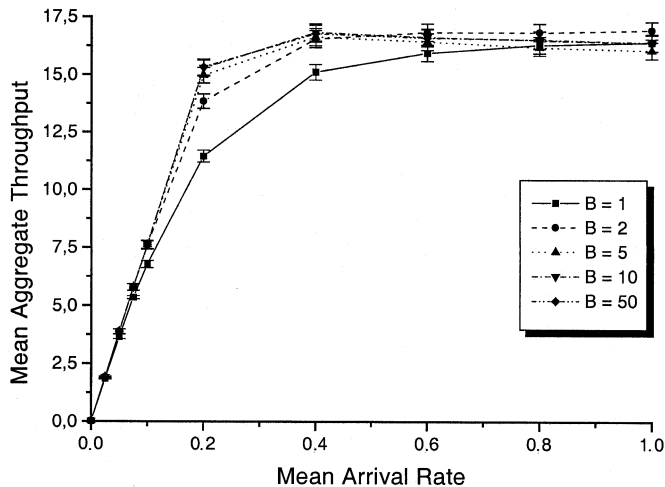


Fig. 23. Mean aggregate throughput (packets per frame) versus mean arrival rate for  $b = 4$ ,  $M = 60$ ,  $R = 8$ ,  $W = 8$ , and different buffer sizes of  $B \in \{1, 2, 5, 10, 50\}$  packets.

window of two cycles, the parameter  $W$  must be equal to eight for  $D = 4$ .

Next, we increase the buffer at each node, i.e., the single-packet buffer is replaced with a buffer that holds up to  $B$  long data packets, where  $B \geq 1$ . (The low-complexity first-in-first-out (FIFO) queuing discipline is considered for our very-high-speed network.) In addition to mean aggregate throughput and mean delay, we consider the relative packet loss (defined as the ratio of dropped packets to the number of generated packets). Fig. 22 illustrates the positive impact of larger buffer sizes on the packet loss. With increasing  $B$ , more arriving data packets can be stored and do not have to be dropped resulting in a decreased packet loss. In addition, each node is less likely to be idle which leads to an increased mean aggregate throughput, as depicted in Fig. 23. However, Fig. 24 shows that the mean delay significantly increases for larger  $B$ .

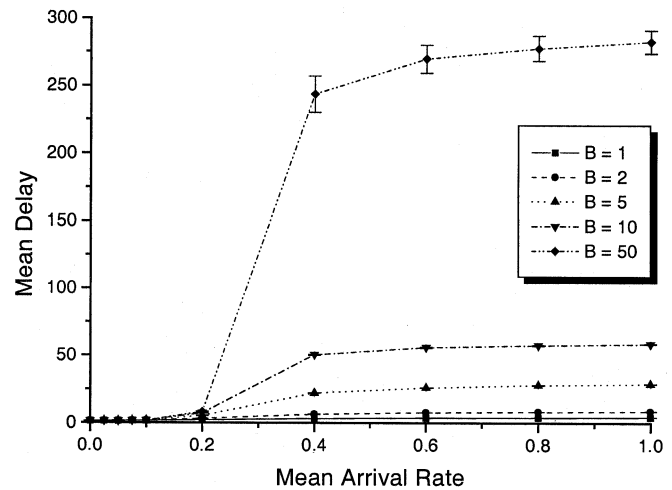


Fig. 24. Mean delay (cycles) versus mean arrival rate for  $b = 4$ ,  $M = 60$ ,  $R = 8$ ,  $W = 8$ , and different buffer sizes of  $B \in \{1, 2, 5, 10, 50\}$  packets.

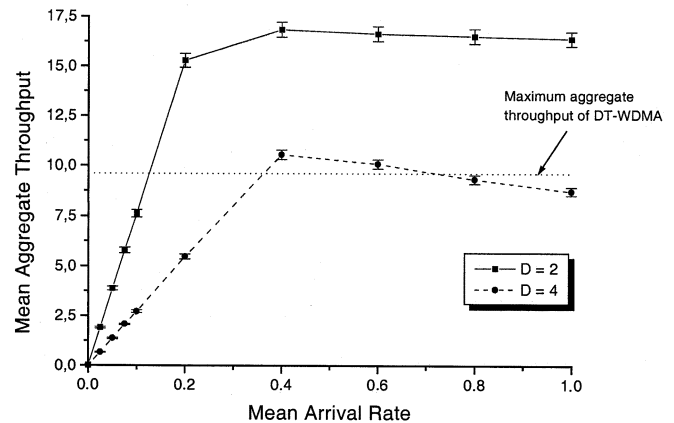


Fig. 25. Mean aggregate throughput (packets per frame) for  $b = 4$ ,  $M = 60$ ,  $W = 8$ , and  $D \in \{2, 4\}$ , compared to maximum aggregate throughput (packets per frame) of DT-WDMA versus mean arrival rate.

This is because in larger buffers there are packets which have to wait for a longer time interval until they are transmitted. Clearly, there is a tradeoff between packet loss and delay. To avoid large delays and provide a reasonable throughput-loss performance we set  $B = 10$  in the following simulations.

We next examine the network performance if the  $N = 200$  nodes are connected by a  $4 \times 4$  AWG instead of a  $2 \times 2$  AWG. Note that changing the AWG degree  $D$  implies also a different  $R$ ,  $F$ , and  $B$ . For a given transceiver tuning range  $\Lambda$ , a larger  $D$  translates into a smaller  $R$  since  $R = \Lambda/D$ . Moreover, we have considered a fixed cycle length of  $D \cdot F = 400$  slots. For  $D = 4$  we get  $F = 100$  slots compared to  $F = 200$  slots for  $D = 2$ . As a consequence, for  $D = 4$ , short data packets are  $K = (F - M) = 40$  slots and long data packets are  $F = 100$  slots long, compared to  $K = 140$  and  $F = 200$  slots for  $D = 2$ , respectively. This results in a decreased mean aggregate throughput, as depicted in Fig. 25 [ignore the dynamic time-wavelength division multiple access (DT-WDMA) curve in this figure for now]. Intuitively, we expect  $D = 4$  to provide a smaller mean delay than  $D = 2$  since for  $D = 4$  each cycle contains twice as many reservation slots as for  $D = 2$  and the



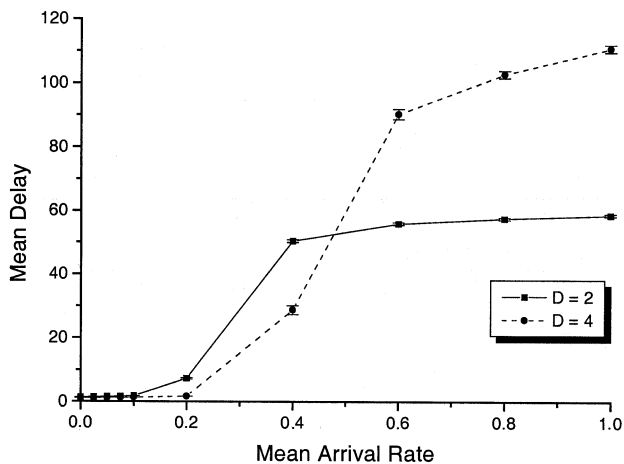


Fig. 26. Mean delay (cycles) versus mean arrival rate for  $b = 4$ ,  $M = 60$ ,  $W = 8$ , and different  $D \in \{2, 4\}$ .

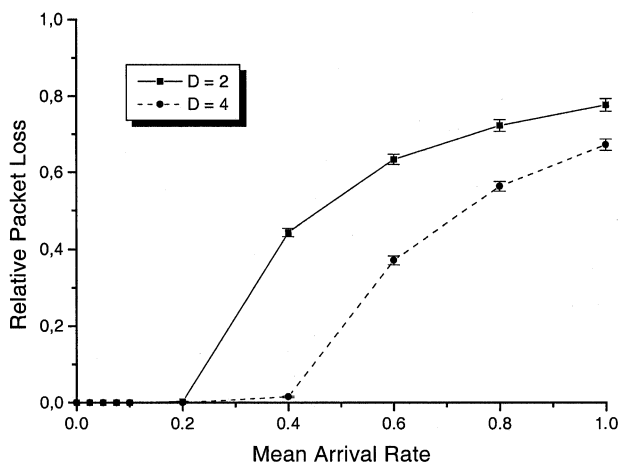


Fig. 27. Relative packet loss versus mean arrival rate for  $b = 4$ ,  $M = 60$ ,  $W = 8$ , and different  $D \in \{2, 4\}$ .

spatial reuse is doubled. We observe from Fig. 26 that this is true only for light-to-medium traffic loads. For a mean arrival rate larger than about 0.45 we observe that  $D = 2$  provides smaller delays. This is due to the fact that each buffer can store up to ten long data packets for  $D = 2$  whereas for  $D = 4$  each buffer is able to store up to 20 long data packets owing to the halved length of long data packets. As a consequence, for  $D = 4$  at medium-to-high loads there are packets which have to wait for a longer time interval until they are transmitted, resulting in an increased mean delay (and a decreased packet loss, as shown in Fig. 27).

### C. Comparison With PSC-Based Network

In this section, we compare our AWG-based network and MAC protocol with a PSC-based single-hop metro WDM network running the so-called *dynamic time-wavelength division multiple access* (DT-WDMA) protocol for resolving packet collisions [8]. We have chosen DT-WDMA since among the MAC protocols specifically designed for multiwavelength single-hop WDM networks based on a PSC, DT-WDMA has probably been the most influential protocol [59]. Moreover, DT-WDMA has

the following properties in common with our protocol, which allow for a reasonably fair comparison.

- For data transmission/reception each node is equipped with *one single* transceiver.
- Each node's receiver is *tunable*. Consequently, receiver collisions can potentially occur and have to be resolved by the access protocol.
- DT-WDMA belongs to the category of *reservation* protocols.
- Resources are dynamically on-demand assigned on a *per-packet* basis.
- Nodes are able to acquire and maintain *global knowledge*.
- Explicit acknowledgments are *not* required.

Next, we briefly describe the network architecture and DT-WDMA (for more details the interested reader is referred to [8]). DT-WDMA is proposed for a metropolitan-sized single-hop WDM network. Each node is equipped with one transceiver fixed tuned to a common control channel. For data each node deploys one transmitter fixed-tuned to a separate wavelength, i.e., each node has its own home channel for transmission, and one tunable receiver. Control information is sent over a dedicated signaling channel. Time is divided into slots (which correspond to frames in our protocol) on each channel, and slots on the control channel are further split into minislots (which correspond to slots in our protocol). Fixed time-division multiple access is used within each slot on the control channel, where one minislot is dedicated to each node. Receivers listen to the control channel and tune to the appropriate channel to receive packets addressed to them. A common distributed arbitration algorithm is used to resolve conflicts when packets from multiple transmitters contend for the same receiver.

Note that in DT-WDMA each node has its own home channel (wavelength) for transmission, i.e., the number of nodes equals the number of wavelengths, while receivers are assumed to be tunable over all these wavelengths. Thus, for large populations, receivers with a large tuning range are required whose large tuning time significantly decreases the channel utilization. In our comparison, we use 16 wavelengths which allows for deploying fast tunable receivers in DT-WDMA (and fast tunable transceivers in our network) whose tuning time is negligible. Furthermore, to compare our above simulation studies with DT-WDMA, we also set the number of nodes to  $N = 200$  in both the AWG- and PSC-based networks. To accommodate 200 nodes in DT-WDMA we let all data wavelengths be equally shared among the nodes while we give each of the 200 nodes its own minislot on the control channel. In our comparison we focus on the throughput. It was shown in [8] that under the same assumptions on packet arrival process (Bernoulli) and traffic pattern (uniform traffic) as made in our preceding simulations, the maximum utilization of each of the 16 wavelength channels is 0.6 in DT-WDMA. This translates into a *maximum* aggregate throughput of  $0.6 \cdot 16 = 9.6$  packets/slot (or 9.6 packets/frame, since one slot in DT-WDMA corresponds to one frame in our protocol).

Fig. 25 compares the mean aggregate throughput of our network versus the mean arrival rate for  $D \in \{2, 4\}$  with the maximum aggregate throughput of DT-WDMA. We observe that for  $D = 4$ , the mean aggregate throughput of our network

is approximately equal to the maximum aggregate throughput of DT-WDMA at medium-to-high traffic loads. However, for  $D = 2$ , our network clearly outperforms DT-WDMA. For a wide range of arrival rates, our proposed network provides a mean aggregate throughput that is about 70% larger than the maximum aggregate throughput of DT-WDMA. To interpret this result, recall that 16 wavelengths are deployed in both PSC- and AWG-based networks. However, due to spatial wavelength reuse, the  $2 \times 2$  AWG provides twice as many communication channels than the PSC leading to the observed larger aggregate throughput.

## VII. CONCLUSION

We have analyzed the photonic switching of variable-size packets with spatial wavelength reuse in an AWG-based metro WDM network. We have obtained computationally efficient and accurate expressions for the throughput and delay in the network. Based on our analytical results, we have conducted extensive numerical investigations of the performance characteristics of the network. We have also conducted extensive simulations to verify the accuracy of the analytical results. Our numerical results indicate that the AWG-based single-hop network, originally proposed in [7], can efficiently transport packets of different sizes. We found that spatial wavelength reuse is crucial for efficient photonic packet switching. Spatial wavelength reuse significantly increases the throughput while dramatically reducing the delay. We have also demonstrated that the spatial wavelength reuse in the AWG-based network gives a mean aggregate throughput that is roughly 70% larger than the maximum aggregate throughput of a PSC-based network running the DT-WDMA protocol.

In our ongoing work we are studying the optimal tradeoffs of the network parameters, e.g., the AWG degree that maximizes the throughput (and minimizes the delay) for a given number of nodes (and traffic load). We are also studying multicasting in the AWG-based metro single-hop network.

## APPENDIX A

### REFINED APPROXIMATION OF $P(Z = k)$

In this appendix, we derive a refined approximation for the distribution of  $Z$ , the number of successful control packets destined to a given AWG output port in a given frame. This refined approximation does not approximate the binomial distributions of the random variables  $Y_i^n$  [see (4)] and  $Y_i^o$  [see (7)] with Poisson distributions. From (13), we note that  $P(X_i = 1) = P(Y_i = 1)$ . Recalling that  $Y_i = Y_i^n + Y_i^o$ , we have

$$P(Y_i = 1) = P(Y_i^n = 1, Y_i^o = 0) + P(Y_i^n = 0, Y_i^o = 1). \quad (52)$$

By the independence of the  $Y_i^n$ 's and  $Y_i^o$ 's we have

$$P(Y_i = 1) = P(Y_i^n = 1)P(Y_i^o = 0) + P(Y_i^n = 0)P(Y_i^o = 1). \quad (53)$$

Hence, with (4) and (7) we obtain

$$P(Y_i = 1) = \eta \frac{\sigma}{M} \left(1 - \frac{\sigma}{M}\right)^{\eta-1} \left(1 - \frac{p}{M}\right)^{S-\eta} + \left(1 - \frac{\sigma}{M}\right)^{\eta} (S - \eta) \frac{p}{M} \left(1 - \frac{p}{M}\right)^{S-\eta-1} \quad (54)$$

TABLE I  
MEAN AGGREGATE THROUGHPUT  $TH_{\text{net}}$  OBTAINED WITH UNREFINED APPROXIMATION, REFINED APPROXIMATION, AND SIMULATION FOR DEFAULT NETWORK PARAMETERS

$\sigma$	0.02	0.04	0.1	0.2	0.5	1.0
$TH_{\text{net, unrefined}}$	0.886	1.77	4.29	7.32	8.45	8.10
$TH_{\text{net, refined}}$	0.888	1.77	4.29	7.37	8.52	8.16
$TH_{\text{net, simul.}}$	0.883	1.77	4.29	7.32	8.48	8.14

$$= \frac{1}{M} \left(1 - \frac{\sigma}{M}\right)^{\eta-1} \left(1 - \frac{p}{M}\right)^{S-\eta-1} \cdot \left[ \eta \sigma \left(1 - \frac{p}{M}\right) + p(S - \eta) \left(1 - \frac{\sigma}{M}\right) \right] \quad (55)$$

$$\approx \frac{S}{M} \left(1 - \frac{\sigma}{M}\right)^{\nu S-1} \left(1 - \frac{p}{M}\right)^{S(1-\nu)-1} \cdot \left[ \nu \sigma \left(1 - \frac{p}{M}\right) + p(1 - \nu) \left(1 - \frac{\sigma}{M}\right) \right] \quad (56)$$

$$=: \kappa \quad (57)$$

where (56) follows by approximating  $\eta/S$  by its expectation  $\nu$ . Thus, the refined approximation of the distribution of  $Z$  is given by

$$P(Z = k) = \binom{M}{k} \left(\frac{\kappa}{D}\right)^k \left(1 - \frac{\kappa}{D}\right)^{M-k}, \quad k = 0, 1, \dots, M. \quad (58)$$

### A. Numerical Evaluation of Refined Approximation for

$P(Z = k)$

In this subsection, we evaluate the refined approximation for  $P(Z = k)$  and compare it with the approximation (15), which we henceforth refer to as ‘‘unrefined.’’ First, note that in the unrefined approximation, the binomial distribution  $\text{BIN}(\eta, \sigma/M)$  is approximated by the Poisson distribution with parameter  $\eta\sigma/M$  [and the  $\text{BIN}((S - \eta), p/M)$  distribution is approximated by the Poisson distribution with parameter  $(S - \eta)p/M$ ]. Clearly, this approximation is accurate when i)  $S$  is large, ii)  $\sigma$  is small (and also  $p$ ), and iii)  $M$  is large. We also note that the evaluation of the refined approximation is computationally slightly more demanding, as it involves the evaluation of the expression in (56) as compared to the single exponential term in (15).

In Table I, we compare the mean aggregate throughput  $TH_{\text{net}}$  obtained with unrefined approximation, refined approximation, and simulation for the typical network parameters chosen as default parameters in Section V [i.e.,  $R = 2$ ,  $q = 0.25$ ,  $M = 30$ ,  $D = 4$ ,  $N = 2000$  (and, thus,  $S = 50$ ),  $p = 0.8$ ,  $F = 200$ , and  $K = 170$ ]. We observe that i) the throughput obtained by both analytical approximations almost coincides, and ii) the analytical results match very well with the simulation results. (Similar observations hold for the average delay.) In fact, we found in our extensive numerical investigations that these two observations hold for all parameter values considered in Section V. In particular, both analytical approximations essentially coincide and match very well with the simulations for  $S$  values as small as 10. We also found that both analytical approximations essentially coincide and match very well with the simulations for  $M$  values as small as 15. Table II compares the throughput obtained

TABLE II  
MEAN AGGREGATE THROUGHPUT  $TH_{\text{net}}$  OBTAINED WITH UNREFINED APPROXIMATION, REFINED APPROXIMATION, AND SIMULATION FOR  $M = 8$

$\sigma$	0.02	0.04	0.1	0.2	0.5	1.0
$TH_{\text{net, unrefined}}$	0.966	1.90	0.331	0.285	0.266	0.260
$TH_{\text{net, refined}}$	0.966	1.90	0.274	0.241	0.226	0.221
$TH_{\text{net, simul.}}$	0.966	1.89	0.272	0.242	0.224	0.223

with unrefined approximation, refined approximation, and simulation for  $M = 8$  and  $K = 192$ , and all other network parameters at their default values. For this small  $M$  value, we observe that unrefined approximation and refined approximation give identical throughput for small  $\sigma$  values ( $\sigma = 0.02, 0.04$ ), but differ for larger  $\sigma$  values. We also observe that the refined approximation matches the simulation results very well. In summary, we find that the unrefined approximation gives accurate results for a wide range of network parameters. The refined approximation is more accurate when  $M$  is small and  $\sigma$  is large, at the expense of a slightly more demanding computation. However, small  $M$  values typically give poor network performance, as is illustrated in Fig. 9, and may therefore not be desirable in practice.

## APPENDIX B

### PROOF OF INDEPENDENCE OF $Y_i^n$ 'S AND $Y_i^o$ 'S

Let  $\eta = \eta(S)$  be a sequence of positive integers satisfying  $\eta(S) \leq S$  and  $\lim_{S \rightarrow \infty} \eta(S) = \infty$ . Let  $Y_i, i = 1, \dots, M$ , be the number of nodes (among all  $\eta$  nodes) that send a control packet in slot  $i$ .

*Claim:* For all nonnegative integers  $k_1, \dots, k_M$ , we have

$$\lim_{S \rightarrow \infty} \frac{P(Y_1 = k_1, \dots, Y_M = k_M)}{\prod_{i=1}^M \frac{e^{-\eta(\tilde{\alpha}/S)} (\frac{\eta \tilde{\alpha}}{S})^{k_i}}{k_i!}} = 1 \quad (59)$$

i.e., the joint distribution of the random variables  $Y_1, \dots, Y_M$  approaches a product of  $M$  Poisson distributions with identical parameters  $\eta \tilde{\alpha}/S$ . In particular,  $Y_1, \dots, Y_M$  are asymptotically independent as  $S \rightarrow \infty$ .

*Proof:* Fix nonnegative integers  $k_1, \dots, k_M$  and assume that  $\eta(S) \geq k_1 + \dots + k_M$ . The distribution of the random vector  $(Y_1, Y_2, \dots, Y_M, \eta - Y_1 - \dots - Y_M)$  is multinomial with parameters  $(\eta, \varrho, \dots, \varrho, 1 - \varrho M)$ , where  $\varrho = \tilde{\alpha}/S (= \sigma/M)$  (we assume that  $S > \tilde{\alpha}M$ , so that  $1 - \varrho M > 0$ ). Therefore

$$\begin{aligned} P(Y_1 = k_1, \dots, Y_M = k_M) &= \frac{\eta!}{k_1! \dots k_M! (\eta - k_1 - \dots - k_M)!} \\ &\cdot \varrho^{k_1 + \dots + k_M} (1 - \varrho M)^{\eta - k_1 - \dots - k_M} \end{aligned} \quad (60)$$

$$\begin{aligned} &= \frac{\varrho^{k_1 + \dots + k_M}}{k_1! \dots k_M!} \frac{1}{(1 - \varrho M)^{k_1 + \dots + k_M}} e^{\eta \ln(1 - \varrho M)} \\ &\cdot \prod_{j=0}^{k_1 + \dots + k_M - 1} (\eta - j). \end{aligned} \quad (61)$$

Hence

$$\begin{aligned} P(Y_1 = k_1, \dots, Y_M = k_M) &= \frac{1}{(1 - \varrho M)^{k_1 + \dots + k_M}} \\ &\cdot \prod_{i=1}^M \frac{e^{-\eta \varrho} (\eta \varrho)^{k_i}}{k_i!} \\ &\cdot e^{\eta(\ln(1 - \varrho M) + \varrho M)} \prod_{j=0}^{k_1 + \dots + k_M - 1} \frac{\eta - j}{\eta}. \end{aligned} \quad (62)$$

Now note that

A)

$$\lim_{S \rightarrow \infty} \frac{1}{(1 - \varrho M)^{k_1 + \dots + k_M}} = 1 \quad (63)$$

since  $\lim_{S \rightarrow \infty} \varrho(S) = \lim_{S \rightarrow \infty} (\tilde{\alpha}/S) = 0$ .

B)

$$e^{\eta(\ln(1 - \varrho M) + \varrho M)} = e^{\eta(-\varrho M + o(1/S^2) + \varrho M)} = e^{\eta o(1/S^2)} \xrightarrow{S \rightarrow \infty} 1 \quad (64)$$

because  $\eta \leq S$ .

C) Since  $\eta(S) \rightarrow \infty$  we have  $((\eta - j)/\eta) \rightarrow 1$  and therefore

$$\lim_{S \rightarrow \infty} \prod_{j=0}^{k_1 + \dots + k_M - 1} \frac{\eta(S) - j}{\eta(S)} = 1. \quad (65)$$

This proves the claim.

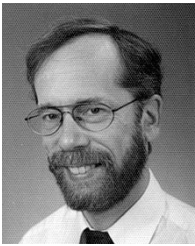
## ACKNOWLEDGMENT

The authors are grateful to Prof. H. Mittelman of Arizona State University for providing them with valuable background on numerical analysis.

## REFERENCES

- [1] C. Awduche and Y. Rekhter, "Multiprotocol lambda switching: Combining MPLS traffic engineering control with optical crossconnects," *IEEE Commun. Mag.*, vol. 39, no. 3, pp. 111–116, Mar. 2001.
- [2] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP-over-WDM networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2062–2071, Oct. 2000.
- [3] D. K. Hunter *et al.*, "WASPNET: A wavelength switched packet network," *IEEE Commun. Mag.*, vol. 37, pp. 120–129, Mar. 1999.
- [4] A. Jourdan *et al.*, "The perspective of optical packet switching in IP-dominant backbone and metropolitan networks," *IEEE Commun. Mag.*, vol. 39, pp. 136–141, Mar. 2001.
- [5] M. J. O'Mahony, D. Simeonidou, D. K. Hunter, and A. Tzanakaki, "The application of optical packet switching in future communication networks," *IEEE Commun. Mag.*, vol. 39, pp. 128–135, Mar. 2001.
- [6] B. Mukherjee, "WDM optical communication networks: Progress and challenges," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1810–1824, Oct. 2000.
- [7] M. Maier, M. Reisslein, and A. Wolisz, "A hybrid MAC protocol for a metro WDM network using multiple free spectral ranges of an arrayed-waveguide grating," *Computer Networks*, vol. 41, no. 4, pp. 407–433, Mar. 2003. A shorter version appeared under the title "High performance switchless WDM network using multiple free spectral ranges of an arrayed-waveguide grating," in *Proc. SPIE Terabit Optical Networking: Architecture, Control, and Management*, Boston, MA, Nov. 2000, pp. 101–112, and won the Best Paper Award of the conference.
- [8] M.-S. Chen, N. R. Dono, and R. Ramaswami, "A media-access protocol for packet-switched wavelength division multiaccess metropolitan area networks," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 1048–1057, Aug. 1990.
- [9] R. Chipalkatti, Z. Zhang, and A. S. Acampora, "Protocols for optical star-coupler network using WDM: Performance and complexity study," *IEEE J. Select. Areas Commun.*, vol. 11, pp. 579–589, May 1993.
- [10] A. Ganz and Y. Gao, "Time-wavelength assignment algorithms for high performance WDM star based systems," *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 1827–1836, Feb./Mar./Apr. 1994.

- [11] I. M. I. Habbab, M. Kavehrad, and C.-E. W. Sundberg, "Protocols for very high-speed optical fiber local area networks using a passive star topology," *J. Lightwave Technol.*, vol. LT-5, pp. 1782–1794, Dec. 1987.
- [12] M. J. Karol and B. Glance, "A collision-avoidance WDM optical star network," *Computer Networks ISDN Syst.*, vol. 26, pp. 931–943, Mar. 1994.
- [13] H. C. Lin and C. H. Wan, "A hybrid multicast scheduling algorithm for single-hop WDM networks," *J. Lightwave Technol.*, vol. 19, pp. 1654–1664, Nov. 2001.
- [14] J. Lu and L. Kleinrock, "A wavelength division multiple access protocol for high-speed local area networks with a passive star topology," *Perform. Eval.*, vol. 16, no. 1–3, pp. 223–239, Nov. 1992.
- [15] N. Mehravari, "Performance and protocol improvements for very high speed optical fiber local area networks using a passive star topology," *J. Lightwave Technol.*, vol. 8, pp. 520–530, Apr. 1990.
- [16] Y. Ofek and M. Sidi, "Design and analysis of hybrid access control to an optical star using WDM," in *Proc. IEEE INFOCOM '91*, Bal Harbour, FL, May 1991, pp. 20–31.
- [17] K. M. Sivalingam, "Design and analysis of a media access protocol for star coupled WDM networks with TT-TR architecture," in *Optical WDM Networks—Principles and Practice*, K. M. Sivalingam and S. Subramaniam, Eds. Norwell, MA: Kluwer, 2000, ch. 9.
- [18] G. N. M. Sudhakar, N. D. Georganas, and M. Kavehrad, "Slotted Aloha and Reservation Aloha protocols for very high-speed optical fiber local area networks using passive star topology," *J. Lightwave Technol.*, vol. 9, pp. 1411–1422, Oct. 1991.
- [19] B. Mukherjee, "WDM-based local lightwave networks part I: Single-hop systems," *IEEE Network Mag.*, vol. 6, pp. 12–27, May 1992.
- [20] M. Maier, M. Reisslein, and A. Wolisz, "Toward efficient packet switching metro WDM networks," *Opt. Networks Mag.*, vol. 3, no. 6, pp. 44–62, Nov./Dec. 2002.
- [21] C. J. Chae, H. Park, and Y. Park, "Hybrid optical star coupler suitable for wavelength reuse," *IEEE Photon. Technol. Lett.*, vol. 10, pp. 279–281, Feb. 1998.
- [22] B. Kannan, S. Fotedar, and M. Gerla, "A two level optical star WDM metropolitan area network," in *Proc. IEEE GLOBECOM '94 Conf.*, San Francisco, CA, Nov. 1994.
- [23] M. W. Janoska and T. D. Todd, "Coupled reservation protocols for hierarchical single-hop photonic networks," *Proc. Inst. Elec. Eng.—Commun.*, vol. 144, no. 4, pp. 247–255, Aug. 1997.
- [24] P. E. Green, L. A. Coldren, K. M. Johnson, J. G. Lewis, C. M. Miller, J. F. Morrison, R. Olshansky, R. Ramaswami, and E. H. Smithand, "All-optical packet-switched metropolitan-area network proposal," *J. Lightwave Technol.*, vol. 11, pp. 754–769, May/June 1993.
- [25] E. Hall, J. Kravitz, R. Ramaswami, M. Halvorson, S. Tenbrink, and R. Thomsen, "The Rainbow-II gigabit optical network," *IEEE J. Select. Areas Commun.*, vol. 14, no. 5, pp. 814–823, June 1996.
- [26] S. Yao, S. J. B. Yoo, and B. Mukherjee, "All-optical packet switching for metropolitan area networks: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 39, pp. 142–148, Mar. 2001.
- [27] D. Stoll, P. Leisching, H. Bock, and A. Richter, "Metropolitan DWDM: A dynamically configurable ring for the KomNet field trial in Berlin," *IEEE Commun. Mag.*, vol. 39, pp. 106–113, Feb. 2001.
- [28] W. Cho and B. Mukherjee, "Design of MAC protocols for DWDM-based metropolitan-area optical ring networks," in *Proc. IEEE GLOBECOM 2001 Conf.*, San Antonio, TX, Nov. 2001.
- [29] C. S. Jelger and M. H. Elmirghani, "A simple MAC protocol for WDM metropolitan access ring networks," in *Proc. IEEE GLOBECOM 2001 Conf.*, San Antonio, TX, Nov. 2001.
- [30] K. Bengi and H. R. van As, "Efficient QoS support in a slotted multihop WDM metro ring," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 216–227, Jan. 2002.
- [31] M. D. Feuer, S. L. Woodward, C. F. Lam, and M. L. Boroditsky, "Upgradable metro networks using frequency-cyclic optical add/drop," in *OFC 2001 Tech. Dig.*, Anaheim, CA, Mar. 2001, paper WBB5.
- [32] F. Ruehl and T. Anderson, "Cost-effective metro WDM network architectures," in *OFC 2001 Tech. Dig.*, Anaheim, CA, Mar. 2001, paper WL1.
- [33] D. Banerjee, J. Frank, and B. Mukherjee, "Passive optical network architecture based on waveguide grating routers," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1040–1050, Sept. 1998.
- [34] B. Glance, I. P. Kaminow, and R. W. Wilson, "Applications of the integrated waveguide grating router," *J. Lightwave Technol.*, vol. 12, pp. 957–962, June 1994.
- [35] B. Glance and M. Karol, "Large capacity multiaccess optical packet network," *IEEE Photon. Technol. Lett.*, vol. 6, pp. 872–875, July 1994.
- [36] A. M. Hill *et al.*, "A scalable and switchless optical network structure, employing a single  $32 \times 32$  free-space grating multiplexer," *IEEE Photon. Technol. Lett.*, vol. 8, pp. 569–571, Apr. 1996.
- [37] D. K. Jung, S. K. Shin, C. H. Lee, and Y. C. Chung, "Wavelength-division-multiplexed passive optical network based on spectrum-slicing techniques," *IEEE Photon. Technol. Lett.*, vol. 10, pp. 1334–1336, Sept. 1998.
- [38] M. J. Spencer and M. A. Summerfield, "WRAP: A medium access control protocol for wavelength-routed passive optical networks," *J. Lightwave Technol.*, vol. 18, pp. 1657–1676, Dec. 2000.
- [39] K. Bengi, "Performance of single-hop WDM LAN's supporting real-time services," *Photonic Network Commun.*, vol. 1, no. 4, pp. 287–301, Dec. 1999.
- [40] A. M. Hill, M. Brierley, R. M. Percival, R. Wyatt, D. Pitcher, K. M. Ibrahim Pati, I. Hall, and J.-P. Laude, "Multi-star wavelength-router network and its protection strategy," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1134–1145, Sept. 1998.
- [41] M. C. Chia, D. K. Hunter, I. Andonovic, P. Ball, I. Wright, S. P. Ferguson, K. M. Guild, and M. J. O. Mahony, "Packet loss and delay performance of feedback and feed-forward arrayed-waveguide gratings-based optical packet switches with WDM inputs outputs," *J. Lightwave Technol.*, vol. 19, pp. 1241–1254, Sept. 2001.
- [42] K. V. Shrikhande *et al.*, "HORNET: A packet-over-WDM multiple access metropolitan area ring network," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2004–2016, Oct. 2000.
- [43] I. M. White *et al.*, "Architecture and protocols for HORNET: A novel packet-over-WDM multiple-access MAN," in *Proc. IEEE GLOBECOM 2000 Conf.*, San Francisco, CA, Nov./Dec. 2000.
- [44] N. P. Caponio, A. M. Hill, F. Neri, and R. Sabella, "Single-layer optical platform based on WDM/TDM multiple access for large-scale 'switchless' networks," *European Trans. Telecommun.*, vol. 11, no. 1, pp. 73–82, Jan./Feb. 2000.
- [45] A. Okada *et al.*, "All-optical packet routing by an out-of-band optical label and wavelength conversion in a full-mesh network based on a cyclic-frequency AWG," in *Proc. OFC 2001 Tech. Dig.*, Anaheim, CA, Mar. 2001, paper ThG5.
- [46] K. Kato *et al.*, "10-Tbps full-mesh WDM network based on a cyclic-frequency arrayed-waveguide grating router," in *Proc. ECOC '00*, vol. 1, Munich, Germany, Sept. 2000, pp. 105–107.
- [47] M. Maier, M. Scheutzow, M. Reisslein, and A. Wolisz, "Wavelength reuse for efficient transport of variable-size packets in a metro WDM network," in *Proc. IEEE INFOCOM '02*, New York, June 2002, pp. 1432–1441.
- [48] C. Dragone, "Optimum design of a planar array of tapered waveguides," *J. Opt. Soc. Amer.*, vol. 7, pp. 2081–2093, Nov. 1990.
- [49] C. Dragone, C. A. Edwards, and R. C. Kistler, "Integrated optics  $N \times N$  multiplexer on silicon," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 896–899, Oct. 1991.
- [50] P. Bernasconi *et al.*, "Large  $n \times n$  waveguide grating routers," *J. Lightwave Technol.*, vol. 18, no. 7, pp. 985–991, July 2000.
- [51] Y. Hibino, "An array of photonic filtering advantages," *Circuits and Devices*, pp. 23–27, Nov. 2000.
- [52] Y. Tachikawa, Y. Inoue, M. Ishii, and T. Nozawa, "Arrayed waveguide grating multiplexer with loop-back optical paths and its applications," *J. Lightwave Technol.*, vol. 14, pp. 977–984, June 1996.
- [53] K. A. McGreer, "Arrayed waveguide gratings for wavelength routing," *IEEE Commun. Mag.*, vol. 36, pp. 62–68, Dec. 1998.
- [54] N. Keil *et al.*, "Athermal polarization-independent all-polymer arrayed waveguide grating (AWG) multi/demultiplexer," in *OFC 2001 Tech. Dig.*, Anaheim, CA, Mar. 2001, paper PD7.
- [55] L. Giehmann, A. Gladisch, N. Hanik, and J. Rudolph, "The application of code division multiple access for transport overhead information in transparent optical networks," in *OFC 1998 Tech. Dig.*, San Jose, CA, Feb. 1998, paper WM42, pp. 228–229.
- [56] L. Giehmann, A. Gladisch, and J. Rudolph, "Field trial of OAM-signal transport capabilities with a 10 Mchip/s LED-direct sequence spread spectrum system suited for OAM-signal-transport in transparent optical WDM-networks," in *Proc. OFC '99*, 1999, paper TuR3, pp. 252–254.
- [57] V. K. Jain and G. DeMarchis, "Performance evaluation of optical code division multiple access networks," *J. Opt. Commun.*, vol. 21, no. 3, pp. 110–115, 2000.
- [58] M. Scheutzow, M. Maier, M. Reisslein, and A. Wolisz. (2003, Feb.) Wavelength reuse for efficient packet-switched transport in an AWG-based metro WDM network. Tech. Rep., Arizona State Univ., Dept. Elec. Eng. [Online]. Available: <http://www.eas.asu.edu/mre>
- [59] R. Ramaswami and K. N. Sivarajan, *Optical Networks—A Practical Perspective*. San Mateo, CA: Morgan Kaufmann, 1998.



**Michael Scheutzow** received the Diploma degree in mathematics from the Johann-Wolfgang-Goethe University, Frankfurt/Main, Germany, in 1979 and the Ph.D. degree in mathematics (*magna cum laude*) from the University of Kaiserslautern, Kaiserslautern, Germany, in 1983. He received the Habilitation in Mathematics also from the University of Kaiserslautern in 1988.

From 1988 through 1990, he was Project Leader at Tecmath GmbH, Kaiserslautern. Since 1990, he has been Professor of Stochastics in the Department of Mathematics at the Technical University of Berlin, Berlin, Germany. From 1997 through 1999, he was Associate Chair of the department. He has visited the University of Carbondale, Carbondale, IL; Rutgers—The State University of New Jersey, New Brunswick; University of Rochester, Rochester, NY; Warwick University, U.K.; and the MSRI at the University of California at Berkeley.



**Martin Maier** received the Dipl.-Ing. and the Dr.-Ing. degrees (both with distinction) in electrical engineering from the Technical University of Berlin, Berlin, Germany, in 1998 and 2003, respectively.

He is currently a Visiting Scholar at the Massachusetts Institute of Technology (MIT), Cambridge. He also participates in the national research project TransiNet. His research interests include switching/routing techniques, protection, architectures, and protocols for optical WDM networks.

Dr. Maier was a recipient of the two-year Deutsche Telekom scholarship from June 1999 through May 2001. He is also a corecipient of the best paper award presented at the SPIE Photonics East 2000—Terabit Optical Networking Conference.



**Martin Reisslein** (A'96–S'97–M'98) received the Dipl.-Ing. (FH) degree from the Fachhochschule Dieburg, Dieburg, Germany, in 1994 and the M.S.E. degree from the University of Pennsylvania, Philadelphia, in 1996, both in electrical engineering. He received the Ph.D. degree in systems engineering also from the University of Pennsylvania in 1998.

He is an Assistant Professor in the Department of Electrical Engineering at Arizona State University, Tempe. During the academic year 1994–1995, he visited the University of Pennsylvania as a Fulbright Scholar. From July 1998 through October 2000, he was a scientist with the German National Research Center for Information Technology (GMD FOKUS), Berlin, Germany. While in Berlin, he was also teaching courses on performance evaluation and computer networking at the Technical University of Berlin. His research interests are in the areas of Internet Quality of Service, video traffic characterization, wireless networking, and optical networking. He maintains an extensive library of video traces for network performance evaluation, including frame size traces of MPEG-4 and H.263 encoded video at <http://www.eas.asu.edu/trace>.

Dr. Reisslein has served on the Technical Program Committees of IEEE INFOCOM and IEEE GLOBECOM. He is Editor-in-Chief of the *IEEE Communications Surveys and Tutorials*. He is corecipient of the Best Paper Award of the SPIE Photonics East 2000—Terabit Optical Networking Conference.



**Adam Wolisz** (M'94–SM'99) received the Dipl.-Ing degree in control engineering, and the Dr.-Ing. degree (both in computer engineering) as well as the Habilitation in 1972, 1976, and 1983, respectively, all from the Silesian Technical University, Gliwice, Poland.

He was with the Polish Academy of Sciences from 1972 to 1989, where he was leading activities on networking, and, from 1990 to 1993, with the German National Research Center for Information Technology (GMD FOKUS), heading activities on quantitative aspects of high-speed networks and multimedia systems. He is a Chaired Professor of Electrical Engineering and Computer Science at the Technical University of Berlin (TUB), Berlin, Germany, where he is heading the Telecommunication Network Group and the Institute of Telecommunication Systems. For the period from 2001 to 2003, he is serving as Dean of the Faculty of Electrical Engineering and Computer Science at TUB. Simultaneously, he is also a member of the Directorial Board of the Fraunhofer Institute for Open Communication Systems (FhG Fokus). His research interests are in architectures and protocols of communication networks as well as protocol engineering with impact on performance and Quality of Service aspects. Currently, he is working mainly on mobile multimedia communication, with special regard to architectural aspects of network heterogeneity and wireless/mobile Internet access. He has authored two books and authored or coauthored over 100 papers in technical journals and conference proceedings.

Prof. Wolisz is a Member of ITG.