The FT$^\Lambda$–FR$^\Lambda$ AWG Network: A Practical Single-Hop Metro WDM Network for Efficient Uni- and Multicasting

Chun Fan, Stefan Adams, and Martin Reisslein, *Senior Member, IEEE*

Abstract—Single-hop wavelength-division-multiplexed (WDM) networks with a central passive star coupler (PSC), as well as single-hop networks with a central arrayed-waveguide grating (AWG) and a single transceiver at each node, have been extensively studied as solutions for the quickly increasing amounts of unicast and multicast traffic in the metropolitan area. The main bottlenecks of these networks are the lack of spatial wavelength reuse in the studied PSC-based networks and the single transceiver in the studied AWG-based metro WDM networks. This paper describes the development and evaluation of the FT$^\Lambda$–FR$^\Lambda$ AWG network, which is based on a central AWG and has arrays of fixed-tuned transmitters and receivers at each node. Transceiver arrays are a mature technology, making the proposed network practical. In addition, the transmitter arrays allow for high-speed signaling over the AWG while the receiver arrays relieve the receiver bottleneck arising from multicasting in conjunction with spatial wavelength reuse on the AWG. The results from probabilistic analysis and simulation reported here indicate that the FT$^\Lambda$–FR$^\Lambda$ AWG network gives particularly good throughput-delay performance for a mix of unicast and multicast traffic.

Index Terms—Arrayed-waveguide grating (AWG), medium access control, multicast, single-hop network, throughput-delay performance, transceiver array.

I. INTRODUCTION

WITH the quickly increasing speeds in the local access networks (due to Gigabit Ethernet and similar emerging technologies) and the provisioning of very-high capacity backbone wavelength-division-multiplexed (WDM) networks, the metropolitan area networks are becoming a bottleneck—the so-called metro-gap. This is largely due to the current circuit-switched synchronous optical network/synchronous digital hierarchy (SONET/SDH) over WDM metro networks, which carry an increasing amount of bursty data and multimedia traffic inefficiently. This situation is further exacerbated by the placement of content distribution proxies in the metro area and the emergence of peer-to-peer networking paradigms. These developments will further increase the traffic load on metro networks. In addition, there will likely be an increase in the portion of multicast (multidestination) traffic in the metro area due to the applications supported by the proxy servers and peer-to-peer networks, such as multimedia stream distribution, distributed games, teleconferences, and telemedicine. Therefore, there is an urgent need for innovative and practical metro networks [1].

Single-hop WDM networks with their minimum hop distance of one (i.e., no bandwidth devoted to multihop packet forwarding) and inherent transparency have attracted a great deal of attention as solutions for the metropolitan area. Single-hop WDM networks are typically either based on a central passive star coupler (PSC) or a central arrayed-waveguide grating (AWG).[1] Each wavelength on the PSC provides a broadcast channel from a given PSC input port to all output ports. Thus, the number of simultaneous transmissions in a PSC network is limited by the number of available wavelengths. Generally, wavelengths are precious, especially for the cost-sensitive metro area and should be utilized efficiently. For this reason, AWG-based networks have recently begun to attract significant attention. The AWG is a wavelength-routing device that allows for spatial wavelength reuse, i.e., the entire set of wavelengths can be simultaneously applied at each AWG input port without resulting in collisions at the AWG output ports. This spatial wavelength reuse has been demonstrated to significantly improve the network performance for a fixed set of wavelengths compared with PSC-based networks [5], [6].

As detailed in Section I-A, the studied AWG-based metro WDM networks employ a single fast-tunable transmitter and a single fast-tunable receiver (TT–TR) at each network node. While this TT–TR node architecture is conceptually very appealing and has a number of advantages, such as lower electronic complexity, low power consumption, and small footprint, fast-tunable transceivers are generally a less mature technology

C. Fan and M. Reisslein are with the Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-5706 USA (e-mail: chun.fan@asu.edu; reisslein@asu.edu; website: http://www.fulton.asu.edu/~mre).

S. Adams is with the Department of Mathematics, Technical University Berlin, 10623 Berlin, Germany (e-mail: adams@math.tu-berlin.de).

[1]It should be noted that WDM ring networks can also be operated as single-hop networks in the sense that packet transmissions optically bypass the intermediate nodes on their way around the ring to the destination node (see, e.g., [2]–[4]). The focus in this paper is on single-hop star WDM networks where packet transmissions do not traverse any intermediate nodes.

than fixed-tuned transceiver arrays. More specifically, fast-tunable transmitters have just recently been experimentally proven to be feasible in a cost-effective manner [7], while fast-tunable optical filter receivers with acceptable channel crosstalk remain a technical challenge at the photonics level. Overall, arrays of fixed-tuned transmitters and receivers are better understood [8], [9], more mature, more reliable, and commercially available, but also have some drawbacks, such as increased electronic complexity, increased power consumption, and larger footprint. The more complex electronics, however, is generally still more mature and less costly than the tunable transceiver solution. At the medium access control (MAC) protocol level, transceiver arrays have a number of distinct advantages. The transmitter arrays allow for high-speed signaling over the AWG as detailed shortly, in contrast to the low-speed signaling through the spectral slicing of broad-band light sources [5], [6], [10] which suffer from a small bandwidth–distance product. The receiver arrays, on the other hand, relieve the receiver bottleneck caused by multicast traffic, which is transmitted over the large number of wavelength channels obtained from spatial wavelength reuse on the AWG.

This paper describes the development and evaluation of the FT^Λ–FR^Λ AWG network, an AWG based single-hop WDM network with an array of fixed-tuned transmitters and receivers at each network node. The proposed FT^Λ–FR^Λ AWG network is practical due to its mature, commercially available building blocks. As we demonstrate through analysis and simulation, the network efficiently supports unicast and multicast traffic. The FT^Λ–FR^Λ node architecture, aside from being readily deployable, achieves good throughput-delay performance especially for a mix of unicast and multicast traffic.

This paper is organized as follows. The following subsection reviews related work. Section II describes the architecture of the FT^Λ–FR^Λ AWG network and discusses how it supports unicast and multicast traffic. In Section III, the distributed MAC protocol is provided. In Section IV, a probabilistic model to evaluate the throughput-delay performance of the network for a mix of unicast and multicast traffic is developed. This analysis considers an operation of the network with essentially no packet drops, achieved with sufficiently large (electronic) node buffers, and is based on a virtual buffer model of the network. Section V presents numerical throughput-delay results obtained from the proposed analytical model and simulations. In Section VI, the node buffer dimensioning for the network is studied, and it is demonstrated that small node buffers are sufficient to achieve minuscule drop probabilities. The findings are summarized in Section VII.

A. Related Work

Both unicasting (see surveys [11] and [12]) and multicasting (see, e.g., [13]–[26] as well as surveys [27]–[29]) over PSC-based networks have been studied extensively. The studied PSC-based networks include networks with arrays of fixed-tuned receivers (see, e.g., [30]), as well as networks with arrays of fixed-tuned transmitters and receivers (see, e.g., [31]). The key bottleneck in the PSC-based network is the channel resource limitation due to the lack of spatial wavelength reuse.

Recently, the use of the wavelength-routing AWG as the central hub in single-hop networks has received more attention. The spatial wavelength reuse of the AWG overcomes the channel resource limitations of single-hop PSC-based networks. The photonic feasibility aspects of the single-hop WDM networks based on a uniform-loss cyclic-frequency AWG with nodes consisting of individual transceivers as well as transceiver arrays have been demonstrated in [32] and [33]. General design principles for networks based on AWGs are studied, for instance, in [34]–[45].

SONATA [46], [47] is a national-scale network based on an AWG. In SONATA, individual nodes (terminals) are connected to passive optical networks (PONs), which in turn are connected to the AWG. SONATA employs a centralized network controller to arbitrate the access of the terminals to the shared wavelength channels and wavelength converter arrays at the central AWG to balance the load between PON pairs. In contrast, a metropolitan area network is considered in this paper with decentralized MAC in which all network nodes maintain global knowledge and execute the same scheduling algorithm in parallel. The proposed network is completely passive and does not employ any wavelength converters.

Unicasting and multicasting in a single-hop AWG-based metro WDM network with decentralized media access control are also studied in [5] and [6]. The network considered in [5] and [6] employs the TT–TR node architecture, which results in slow signalling and the receiver bottleneck. It is also noted that the analytical performance model of the multicasting in the TT–TR AWG network developed in [6] considers a simplified multicast traffic model in which a multicast packet is destined to all nodes attached to exactly one of the AWG output ports. In contrast, in this paper an analytical performance model is developed for the multicasting in the FT^Λ–FR^Λ AWG network, which considers the more realistic and widely accepted multicast traffic model with randomly uniformly distributed number and location of destinations of a multicast packet.

In this paper, the focus is on the network and MAC protocol design of the FT^Λ–FR^Λ AWG network and its performance evaluation. The protection and survivability aspects of the network are beyond the scope of this paper. It is noted that protection strategies for AWG-based networks have been examined in [48]–[50]. In their ongoing work, the authors are developing similar strategies for the FT^Λ–FR^Λ network.

II. ARCHITECTURE

Our AWG-based network architecture is illustrated in Fig. 1. We consider a cyclic AWG with D input ports and D output ports, whose free spectral range (FSR) is equal to the number of ports (times the channel spacing, to be consistent with the units). There are N nodes in the network. At each AWG input port, an $S \times 1$, $S = N/D$, combiner collects transmissions from the transmitters of S attached nodes. At each AWG output port, a $1 \times S$ splitter equally distributes the signal to S individual fibers that are attached to the receivers of the nodes. We use the notation $N_{i,j}$, $i = 1, 2, \ldots, D$, $j = 1, 2, \ldots, S$ to designate the jth node attached to the ith AWG port. In Fig. 1, $T_{i,j}$ and $R_{i,j}$ correspond to the transmitter array and the receiver array of node $N_{i,j}$, respectively.
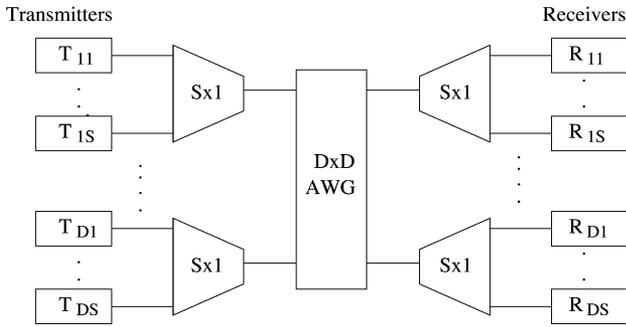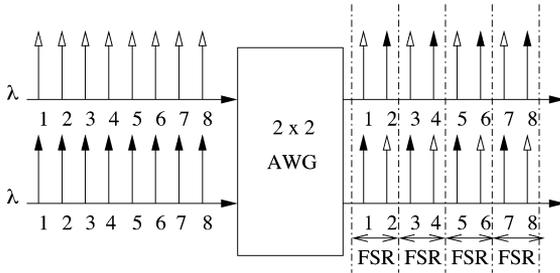
Fig. 1. Network architecture.

Fig. 2. Wavelength routing in 2×2 AWG with $R = 4$ FSRs.

The wavelength-routing property of the AWG is illustrated in Fig. 2 for a 2×2 AWG (i.e., an AWG with degree $D = 2$) with a period of the wavelength response (referred to as the FSR) of $R = 4$. According to the periodic wavelength routing, every Dth wavelength is routed to the same AWG output port. Note that two transmissions on different wavelengths are required to reach both AWG output ports from a given input port. Also note that $\Lambda = D \cdot R$ wavelength channels can be simultaneously used at each of the D AWG ports without resulting in channel collisions. With this "spatial reuse" of wavelength channels, the AWG provides a total of $D \cdot \Lambda$ channels from its D input ports to its D output ports. There are R channels between each input–output port pair.

The node architecture is shown in Fig. 3. Each node is equipped with a transmitter array consisting of Λ fixed-tuned transmitters and a receiver array consisting of Λ fixed-tuned receivers. The optical multiplexer is used to combine multiple transmissions from the node's transmitter array onto the transmission fiber. The optical demultiplexer is used to separate the signal from the receiving fiber to the receiver array. To keep the node structure simple, we consider a single queue at each node, which buffers both the unicast packets and multicast packets generated by the node.

We end this overview of the FT$^\Lambda$–FR$^\Lambda$ AWG network architecture by noting its implications on the transmission of unicast and multicast packets. A unicast packet, that is, a packet that is destined to one destination node, requires one transmission on one of the wavelengths that is routed to the AWG output port to which the destination node is attached.

Now consider a multicast packet, that is, a packet that is destined to two or more destination nodes. If all destination nodes are attached to the same AWG output port, then only one transmission is required on the wavelength routed to that AWG output port. The splitter locally broadcasts the transmission to
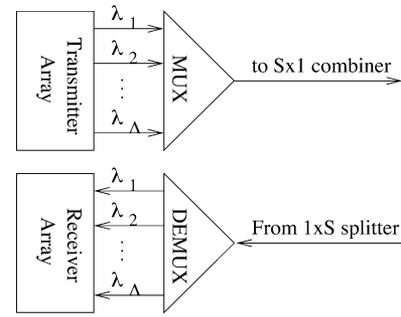

Fig. 3. Detailed node architecture.

all attached nodes, including the intended destination nodes. On the other hand, if the destination nodes of a given multicast packet are attached to different AWG output ports, transmissions on multiple wavelengths routed to the different AWG output ports are required. Note that the FR$^\Lambda$ receiver array at each node eliminates receiver conflicts (receiver collisions) since all Λ wavelength channels connecting a splitter to an AWG output port can be simultaneously received by all nodes attached to the splitter. Hence, all destination nodes of a multicast packet copy transmitted on one of the Λ wavelengths are reached by the copy transmission, which eliminates the need for splitting the fan-out of multicast packet copies. As discussed in the next section in more detail, these multiple transmissions can be conducted in parallel using multiple transmitters in the source node's transmitter array at the same time.

III. MAC PROTOCOL

In this section, we develop a MAC protocol employing pretransmission coordination together with global scheduling to coordinate the access of the nodes to the shared wavelength channels in the FT$^\Lambda$–FR$^\Lambda$ AWG network. This coordination and scheduling are generally recommended strategies for achieving good throughput-delay performance in shared-wavelength single-hop star networks [11]. Time is divided into frames, with each frame consisting of a control phase and a data phase, as illustrated in Fig. 4. The length of each control packet measured in time is one slot. One control packet is generated for each data packet. The control packet contains the address of the destination node for unicast packets or the multicast group address for multicast packets.

We develop two control packet transmission strategies: time-division multiple access (TDMA) and contention similar to slotted Aloha. With either strategy, the periodic wavelength-routing property of the AWG requires a transmitting node to use all of the wavelengths covering at least one FSR in order to reach all of the AWG output ports. The spatial wavelength reuse property also allows nodes attached to different ports of the AWG to use the same set of wavelengths without channel collision.

A. TDMA Control Packet Transmission

The TDMA sequence for control packet transmission in an AWG network with *one* FSR ($R = 1$) is as follows: in the first slot of the control phase, one node from each input port of the AWG, say the first node $N_{d,1}$ at each port $d = 1, 2, \ldots, D$,
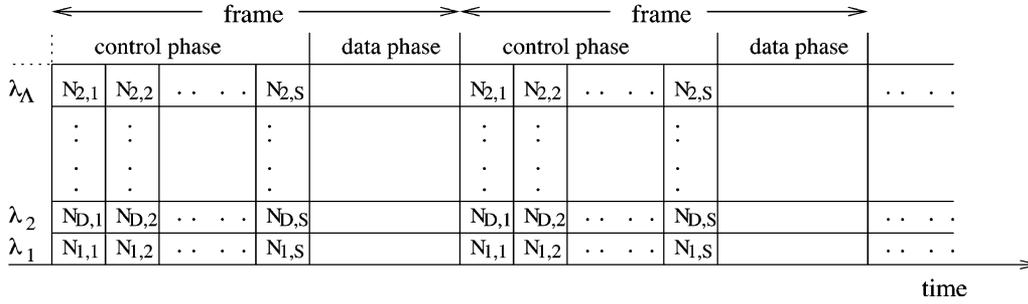
Fig. 4. Frame structure and control packet reception schedule for nodes at AWG output port 1 of network with $R = 1$ FSR.
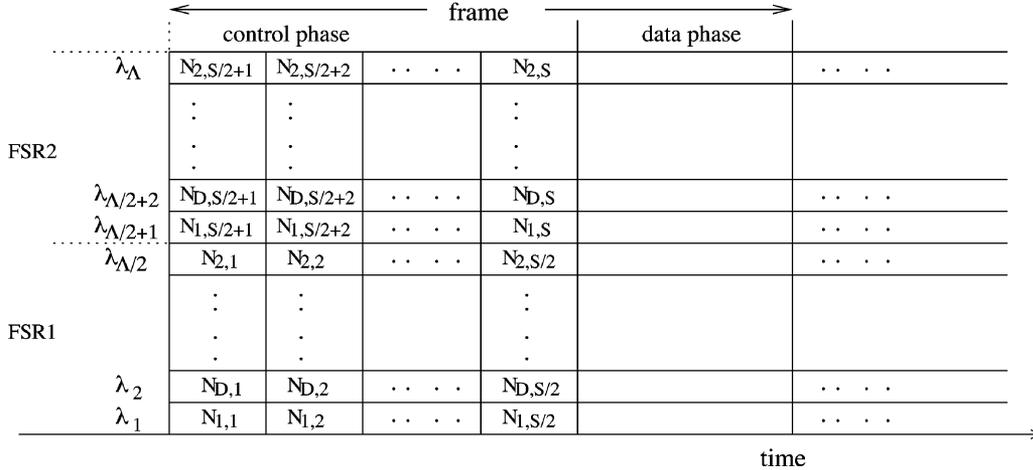
Fig. 5. Frame structure and control packet reception schedule for nodes at AWG output port 1 of network with $R = 2$ FSRs.

transmits its control packet, whereby the data packets in the node's queue are considered in first-come–first-served (FC–FS) manner to ensure low complexity. Each node uses its full array of fixed transmitters for control packet transmission, which are conducted at the full data bit rate of the lasers. In contrast, in a single transceiver AWG network, a broad-band light source in conjunction with control signal spreading and spectral slicing is required to maintain global knowledge of the control packet transmissions at all nodes; this signaling with the broad-band light source results in a significantly smaller signaling bit rate compared with the data bit rate of a laser [5], [6]. In the second slot, another node from each AWG input port, say the second node $N_{d,2}$ at each port $d = 1, 2, \ldots, D$, transmits its control packet. This continues until all of the nodes have transmitted their control packets. Fig. 4 shows the corresponding control packet reception schedule by the receiver array of the nodes at AWG output port 1; the reception schedules for the other output ports are analogous. To understand the reception schedule in Fig. 4, observe that by the periodic wavelength-routing property of the AWG, wavelength 1 from AWG input port 1 is routed to the considered AWG output port 1, wavelength 2 from AWG input port D is routed to output port 1, and so on. Note that the control packets do not need to carry the source address, as the source node address can be inferred from the reception schedule. The control phase is S slots long. (Recall that $S = N/D$ and $\Lambda = D \cdot R$. In the considered case $R = 1$, we have $\Lambda = D$ and thus $S = N/\Lambda$.)

In the case of a network with R FSRs, we split the nodes attached to each AWG port into R subgroups. Each subgroup is given a different FSR for the transmission of the control packets. Thus, we have R nodes from each input port simultaneously transmitting control packets, with each node using all wavelengths in one of the R FSRs. The control packet reception schedule for the nodes at AWG output port 1 of a $R = 2$ FSR network is shown in Fig. 5.

In general, the length of the control phase with the TDMA transmission strategy is S/R slots. Note, however, that $S = N/D$ and $R = \Lambda/D$ results in a constant control phase length of N/Λ slots, independent of the number of FSRs R. In other words, the length of the control phase depends only on the number of nodes N and the number of transceivers Λ at each node. Consequently, in our performance evaluations in Section V, we do not need to explicitly include the control phase when considering scenarios with TDMA control packet transmission with fixed N and Λ. When comparing scenarios with different TDMA control phase lengths N/Λ or control packet contention, we take the different lengths of the control phase into consideration.

B. Control Packet Transmission With Contention

With the contention control packet transmission strategy, the control packets are transmitted similar to slotted Aloha. In a network with $R = 1$ FSR, each node sends the control packet uniformly and randomly in one of the slots of the M, $M \leq N\Lambda$ slot long control phase using its full array of transmitters. In the case of multiple FSRs connecting each input–output port pair, the transmitting node picks from one of the FSRs randomly
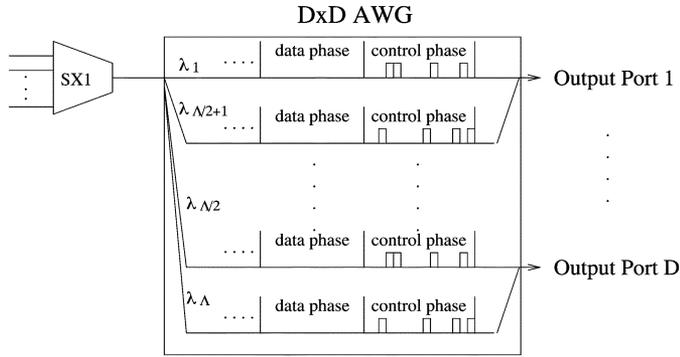
Fig. 6. Control packet contention and frame structure for network with $R = 2$ FSRs; the control phase is M slots long.

and uniformly to transmit the control packet in a uniformly randomly chosen slot on all wavelengths in the selected FSR, as illustrated in Fig. 6 for $R = 2$.

A collision occurs when two or more nodes select the same control slot (in the same FSR). Since the transmitter uses all the wavelength of one full FSR and the receiver arrays cover all of the wavelengths, the transmitting node knows the results of control contention after a delay of the one-way end-to-end propagation delay. The nodes with collided control packets retransmit the control packet in the following frame.

Note that for the control packet contention, the control packet needs to contain the address of the source node in addition to the addresses of the destination nodes.

We also note that in the FT^Λ–FR^Λ AWG network, the R wavelengths (and corresponding receivers) connecting a given AWG input port with a given AWG output port are only shared by the transmissions between nodes attached to these two ports. Thus, the network allows for the development of contention-based MAC protocols where control packets are only sent to the AWG output port(s) with attached receivers. Such protocols would have the advantage that typically fewer lasers are required for a control packet transmission compared with our protocol where control packets are transmitted to all output ports using all lasers in one FSR. One drawback of such protocols would be that the sending node does not necessarily receive a copy of a sent control packet. Thus, explicit acknowledgment would be required to verify whether a control packet collision occurred; these acknowledgments would result in increased protocol complexity and delay. Along the same line, the FT^Λ–FR^Λ AWG network allows for the development of MAC protocols where the data packets contend directly for the R wavelength channels connecting a given AWG input–output port pair without pretransmission coordination. Such uncoordinated data packet contention, however, would tend to result in a significant waste of bandwidth due to data packet collisions [11].

C. Data Packet Scheduling

Once the control packets of a given control phase are received, all nodes execute the same scheduling algorithm. For a unicast packet, as well as for a multicast packet with all destination nodes attached to one AWG output port, a single packet transmission is scheduled. For a multicast packet with destination nodes at multiple AWG output ports, multiple packet (copy)

transmissions are scheduled: one copy is transmitted to each AWG output port with attached multicast destination nodes. A wide variety of algorithms can be employed to schedule the unicast data packets and the multicast data packet copies (corresponding to the received control packets) on the wavelength channels. To avoid a computational bottleneck in the distributed scheduling in the nodes in our very high-speed optical network, the scheduling algorithm must be simple [21]. Therefore, we adopt a FC–FS/first-fit (FF) scheduling policy as follows. All transmission requests (irrespective of whether they are for a unicast packet or a multicast packet copy transmission) are considered on a FC–FS basis and are assigned a wavelength on a FF basis, that is, a wavelength leading to the desired AWG destination port is assigned starting with the lowest FSR in the immediate frame. If there is no free wavelength leading to the desired output port in the immediate frame, then the wavelengths in the subsequent frame are assigned, and so on, up to a prespecified scheduling window. We define the scheduling window as the maximum number of frames that data packets are allowed to be scheduled into the future. We introduce this window to accommodate any limitations on the scheduling memory in practical networks in our protocol; however, we expect in most cases of practical interest the scheduling window to be several tens of frames long and have a relatively minor impact as indicated by our delay results in Section V. If the unicast data packet or some of the multicast data packet copies corresponding to a control packet can not be scheduled within the scheduling window, the control packet fails. The sending node is aware of the failed control packet as it executes the same scheduling algorithm and retransmits the failed control packet in the next frame. (For a failed control packet corresponding to a multicast data packet, the retransmitted control packet requests only the failed multicast packet copy transmissions.)

Note that unfairness among the nodes may arise with the FC–FS scheduling if the control packets are transmitted (and received) in the fixed TDMA sequence. To overcome this problem, the received control packets can be randomly resequenced before the scheduling commences. Control packet contention also ensures fairness since the control packets are transmitted in randomly selected slots. Also note that both randomly resequencing the control packets in the TDMA approach as well as the control packet contention approach ensure fairness with respect to being scheduled within the scheduling window since, with both approaches, *each* transmitted control packet (irrespective of its source node) has the same chance of succeeding or failing in the scheduling.

Note that the data packets are buffered in the electronic domain at each source node which can have quite large memory capacity. An arriving packet that finds the node buffer full is dropped and is indicative of congestion. We leave traffic congestion management to the upper layer protocols.

In Fig. 7, we illustrate the MAC protocol and data packet scheduling with an example for a 2×2 AWG with $R = 2$ FSRs and $S = 4$ nodes attached to each AWG port. The illustration shows the signals transmitted by the transmitters attached (via a combiner) to AWG input port 1, as well as the signals received by the receivers attached to AWG port 1, whereby the time axes at the transmitters and receivers are offset by the one-way

Fig. 7 (transmitter schedule):

		frame (control)	frame (data)		frame (control)	frame (data)	frame (control)	frame (data)	
t r a n s	FSR2 λ_4	$N_{1,3}$ $N_{1,4}$		$\cdots$	$N_{1,3}$ $N_{1,4}$	$N_{1,1}$	$N_{1,3}$ $N_{1,4}$		$\cdots$
	λ_3	m12		$\cdots$					$\cdots$
	FSR1 λ_2	$N_{1,1}$ $N_{1,2}$		$\cdots$	$N_{1,1}$ $N_{1,2}$	$N_{1,2}$	$N_{1,1}$ $N_{1,2}$	$N_{1,4}$	$\cdots$
	λ_1	u2 u2		$\cdots$		$N_{1,4}$			$\cdots$

Fig. 7 (receiver schedule):

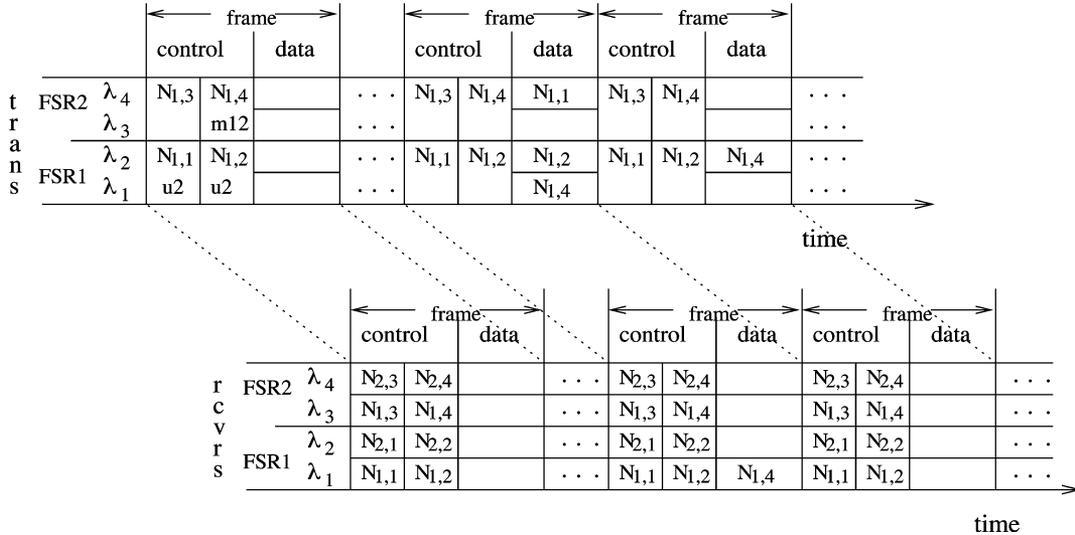|		frame (control)	frame (data)		frame (control)	frame (data)	frame (control)	frame (data)	
r c v r s	FSR2 λ_4	$N_{2,3}$ $N_{2,4}$		$\cdots$	$N_{2,3}$ $N_{2,4}$		$N_{2,3}$ $N_{2,4}$		$\cdots$
	λ_3	$N_{1,3}$ $N_{1,4}$		$\cdots$	$N_{1,3}$ $N_{1,4}$		$N_{1,3}$ $N_{1,4}$		$\cdots$
	FSR1 λ_2	$N_{2,1}$ $N_{2,2}$		$\cdots$	$N_{2,1}$ $N_{2,2}$		$N_{2,1}$ $N_{2,2}$		$\cdots$
	λ_1	$N_{1,1}$ $N_{1,2}$		$\cdots$	$N_{1,1}$ $N_{1,2}$	$N_{1,4}$	$N_{1,1}$ $N_{1,2}$		$\cdots$

Fig. 7. Illustrative example of control packet transmission and data packet scheduling: the four nodes at AWG input port 1 have two unicast packets destined to nodes at AWG output port 2 and one multicast packet destined to nodes at both AWG output ports. The time axes at the transmitters and receivers are offset by the propagation delay.

end-to-end propagation delay. In the illustrative example, the first frame node $N_{1,1}$ has a unicast packet for a node at AWG output port 2, as does node $N_{1,2}$. Node $N_{1,4}$ has a multicast packet destined to nodes at both AWG output ports. The control packets corresponding to these data packets are denoted by u2 for the unicast packets destined to a node at port 2 and m12 for the multicast packet destined to both output ports. These control packets are transmitted according to the TDMA schedule, arrive after the propagation delay at the receivers, and are randomly resequenced so that the control packet from node $N_{1,2}$ is considered first, followed by the control packets from nodes $N_{1,1}$ and $N_{1,4}$. The data packets are scheduled into the data phase coming up next (in the scenario drawn in the figure, there is a time duration of approximately one control slot to complete the computation of the schedule) according to the FC–FS–FF scheduling rule. The unicast packet from $N_{1,2}$ is scheduled on the lowest wavelength routed to port 2, namely λ_2. The unicast packet from $N_{1,1}$ is scheduled on the next lowest wavelength leading to port 2, namely λ_4. Finally, the two copies of the multicast packet from $N_{1,4}$ are scheduled on the earliest available and lowest indexed wavelengths leading to the two ports, which results in the two copies being scheduled in different frames. The copy sent on λ_1 arrives at the receivers at port 1 after the propagation delay.

IV. THROUGHPUT-DELAY ANALYSIS BASED ON VIRTUAL QUEUE MODEL

In this section, we develop a probabilistic virtual queue-based model to evaluate the throughput-delay performance of the $FT^\Lambda - FR^\Lambda$ AWG network. We assume in this model that the nodal buffers are sufficiently large (infinite in the model) such that only a negligible fraction of the packets is dropped. We demonstrate in Section VI that a reasonably small packet buffer at each node is sufficient to achieve packet drop rates of 10^{-2} and less, which in turn implies a correspondingly small modeling error due to the infinite buffer assumption. We also
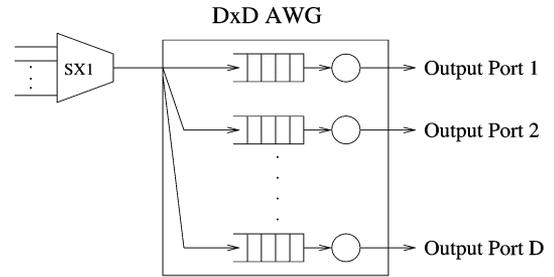
Fig. 8. Queuing model: one *virtual* queue for each AWG input–output port pair. Note that there is no physical buffer at the AWG.

note that throughout, we study the network for stable operation, as detailed in Section IV-D.

A. Overview of Virtual Queue Network Model

We model each AWG input–output port pair as a "virtual" queue. This queue is virtual because there is *no* electronic buffer or optical memory at the AWG. The queue only exists in the electronic memory domain of each node. These virtual queues are illustrated in Fig. 8. The service capacity for a given virtual queue is the number of FSRs R, with each FSR providing a deterministic service rate of one packet per frame.

We consider the following scenario in our modeling of the FT^Λ–FR^Λ AWG network in this section.

- *Bernoulli traffic arrival*: Each node generates a new data packet with probability σ at the beginning of each frame. A given newly generated packet is a unicast packet with probability u and a multicast packet with probability $1-u$. Let $\sigma_u = \sigma \cdot u$ denote the probability that a new unicast packet is generated in a given frame, and let $\sigma_m = \sigma \cdot (1-u)$ denote the probability that a new multicast packet is generated in a given frame.
- *Uniform distribution of traffic*: The destination node(s) of a given unicast (multicast) packet are uniformly distributed over all N nodes, including the sending node for

mathematical convenience. (Our simulations, which do not allow a node to send to itself, indicate that this simplifying assumption has negligible impact.)

- *Uniform multicast size distribution*: We let $\Gamma, 2 \le \Gamma \le N$ represent the maximum number of destination nodes of the multicast packets. The number of destination nodes of a given multicast packet is a random variable γ with $2 \le \gamma \le \Gamma$, which is uniformly distributed, that is, $\gamma \sim U(2, \Gamma)$.
- *Propagation delay*: We initially assume that the propagation delay is negligible. In Section IV-F, we discuss how to incorporate propagation delay in our model.
- *Fixed packet size*: We assume that the data packets are fixed in size. The packet size is such that exactly one data packet fits into the data phase of a given frame.
- *TDMA control packet transmission*: We focus on the TDMA control packet transmission in this paper. The control packet transmission with TDMA and contention are compared in [55].
- *Infinite nodal buffers and scheduling window*.

To model the multiple transmissions of copies of a multicast packet destined to multiple AWG output ports, we place one packet copy into each corresponding virtual queue. Thus, for a multicast packet from a given AWG input port destined to all D AWG output ports, one packet copy is placed in each of the D virtual queues modeling these D input–output port pairs.

B. Definition of Performance Metrics

In our throughput-delay performance evaluation, we consider the following metrics:

- The *multicast throughput* Z_M is defined as the average number of packet transmissions completed per frame in steady state. The transmission of a multicast packet is complete if all copies of the packet have been delivered.
- The *transmitter throughput* Z_T is defined as the average number of packet (copy) transmissions per frame in steady state.
- The *receiver throughput* Z_R is defined as the average number of packets received by their intended destination nodes per frame in steady state. Each intended destination node of a multicast packet copy transmission counts toward the receiver throughput. A given multicast packet copy transmission can result in up to S received packets in case all nodes attached to the splitter are intended destinations.
- The *delay* W_M is the average time in steady state in frames between the following two epochs: 1) the end of the control phase of the frame in which a packet is generated and 2) the beginning of the data phase in which the *last* copy of the packet is transmitted.
- The *copy delay* W_{TR} is defined similar to the delay W_M and is the average time between packet generation and the beginning of the transmission of *any given (arbitrary)* copy of the packet.

Note that when only unicast traffic is considered, $Z_M = Z_T = Z_R$ and $W_M = W_{\text{TR}}$. Also note that all of these perfor-

mance metrics are defined with respect to the frame as elementary time unit. This is convenient as for most of our performance studies we consider a network with fixed number of nodes N and fixed number of transceivers Λ per node. For this network, the length of the TDMA control phase N/Λ is constant, which in conjunction with the fixed data phase (data packet size) results in a constant frame length. Toward the end of our performance evaluation, we will study networks with different N and Λ as well as control packet contention and consequently different frame lengths. For those studies, we will modify the above definitions and use the slot as elementary time unit. In addition, for all experiments using the slot as time unit, we define the delay as the average period between the packet generation (at the beginning of a frame) and the beginning of the packet transmission, which includes the duration of the control phase.

C. Number of Packet Copies

In this section, we evaluate the number of packet copy transmissions required to service a given generated packet. Let Δ be a random variable denoting the number of AWG output ports (virtual queues) that lead to destination nodes of a given generated packet. In other words, Δ denotes the number of packet copies that are placed in different virtual queues for a given generated packet. A single packet copy is transmitted if either 1) the generated packet is a unicast packet (which has probability u) or 2) the generated packet is a multicast packet (which has probability $1 - u$) and all the destination nodes are attached to the same AWG output port. If a multicast has destinations at $l, 2 \le l \le D$, AWG output ports, then l packet copies are generated and one each is placed in the corresponding virtual queue.

To evaluate the number of packet copies required to service a given generated multicast packet, we need to find the number of AWG output ports that have at least one destination node of the packet attached. Toward this end, we model the N nodes attached to the D AWG output ports as an urn containing N balls in D different colors, i.e., there are $S (=N/D)$ balls of color $i, i = 1, \ldots, D$. Suppose the considered multicast packet has $\gamma, 2 \le \gamma \le \Gamma$ destinations. To determine the number of packet copy transmissions, we draw γ balls (each representing a destination node) from the urn without replacement. (An urn model with replacement, which is a simpler, less accurate model of the multicasting, is developed in [51]. In Appendix B we examine the differences between the urn models with and without replacement.) We consider the outcome of the drawing without replacement and study formally the following events: $C_{k_1,\ldots,k_D} =$ "Event that among γ balls drawn without replacement, color 1 occurs k_1 times, color 2 occurs k_2 times, ..., color D occurs k_D times with $k_1 + \cdots + k_D = \gamma$."

The probability of this event is given by the polyhypergeometric distribution [52], which can be obtained from the hypergeometric distribution [53], as follows:

$$P(C_{k_1,\ldots,k_D}) = \frac{\binom{S}{k_1} \cdots \binom{S}{k_D}}{\binom{N}{\gamma}}. \qquad (1)$$

The family of events

$$
C_{k_1,\ldots,k_D}\left(0 \le k_i \le \gamma \wedge S; i = 1,\ldots,D; \sum_{i=1}^{D} k_i = \gamma\right) \quad (2)
$$

forms a complete system of independent events. Thus

$$
P\left(\bigcup_{\substack{0 \le k_i \le \gamma \wedge S; 1 \le i \le D \\ \sum_{i=1}^{D} k_i = \gamma}} \{C_{k_1,\ldots,k_D}\}\right)
$$
$$
= \sum_{\substack{0 \le k_i \le \gamma \wedge S; 1 \le i \le D \\ \sum_{i=1}^{D} k_i = \gamma}} P(C_{k_1,\ldots,k_D}) = 1. \quad (3)
$$

Note that we denote $x \wedge y := \min(x, y)$. In our model, the number Δ of the required packet copy transmissions corresponds to the number of distinct colors among the γ balls drawn without replacement. Toward the evaluation of the distribution of Δ, we define the set of color number vectors

$$
A_\gamma^l = \left\{(k_1,\ldots,k_D) \in \{0,\ldots,S \wedge \gamma\}^D \,\middle|\, \exists k_{i_1},\ldots,k_{i_l}; i_s \right.
$$
$$
\in \{1,\ldots,D\}, 1 \le s \le l \text{ with } k_{i_s} \ge 1 \text{ and}
$$
$$
\left. \sum_{s=1}^{l} k_{i_s} = \gamma; k_r = 0 \text{ for } r \ne i_s, 1 \le r \le D\right\} \quad (4)
$$

for $1 \le l \le D \wedge \gamma$. Intuitively, this is the set of all color number vectors $(k_1,\ldots,k_D)$ such that there are l distinct colors among the drawn γ balls. The probability that the number Δ of required packet copy transmissions for a given multicast packet with γ destinations is l is then given by

$$
P(\Delta = l \,||\, \gamma) = \sum_{(k_1,\ldots,k_D) \in A_\gamma^l} P(C_{k_1,\ldots,k_D}). \quad (5)
$$

Noting that there are $\binom{D}{l}$ ways of choosing l colors out of the D colors (i.e., choosing l destination ports out of all D AWG output ports), we obtain

$$
P(\Delta = l \,|\, \gamma = n) = \binom{D}{l} \sum_{\substack{1 \le k_1,\ldots,k_l \le n \wedge S; \\ \sum_{i=1}^{l} k_i = n}} \frac{\binom{S}{k_1} \cdots \binom{S}{k_l}}{\binom{N}{n}}
$$
$$
(6)
$$

which can be readily computed via recursion, as detailed in Appendix A.

Note that we have calculated in (6) the conditional probability of the event that the number of required packet copies is l given that the generated multicast packet has γ destination nodes, i.e.,

$$
P(\Delta = l; \gamma \text{ dest. nodes}) = P(\Delta = l \,|\, \gamma \text{ dest. nodes})
$$
$$
\cdot P(\gamma \text{ dest. nodes}) \cdot P(\text{multicast}) \quad (7)
$$

with $P(\gamma \text{ dest. nodes}) = (1/(\Gamma-1))$ and $P(\text{multicast}) = 1-u$. As noted previously, a single packet copy is transmitted if either a unicast packet is generated or the generated multicast packet has all $\gamma, 2 \le \gamma \le S \wedge \Gamma$ destination nodes attached to the same AWG output port, i.e.,

$$
P(\Delta = 1) = P(\text{"gen unicast pkt"}) + P(\text{"gen.}
$$
$$
\text{multicast pkt has all dest. at one port"})
$$
$$
= u + \frac{(1-u)}{(\Gamma-1)} \sum_{\gamma=2}^{S \wedge \Gamma} P(\Delta = 1 \,|\, \gamma)
$$
$$
= u + \frac{(1-u)}{(\Gamma-1)} \sum_{\gamma=2}^{S \wedge \Gamma} \frac{D\binom{S}{\gamma}\binom{S}{0} \cdots \binom{S}{0}}{\binom{N}{\gamma}} \quad (8)
$$
$$
= u + \frac{(1-u)D}{\Gamma-1} \sum_{\gamma=2}^{S \wedge \Gamma} \frac{S!(N-\gamma)!}{(S-\gamma)!N!}. \quad (9)
$$

The probability that a given generated packet has destinations at $l, 2 \le l \le D$ AWG output ports, that is, requires l packet copy transmissions, is

$$
P(\Delta = l) = \frac{(1-u)}{\Gamma-1} \sum_{n=2}^{\Gamma} P(\Delta = l \,|\, \gamma = n). \quad (10)
$$

We obtain the expected number of required packet copy transmissions as

$$
E[\Delta] = P(\Delta = 1) + \sum_{l=2}^{D} l P(\Delta = l)
$$
$$
= u + \frac{(1-u)D}{\Gamma-1} \sum_{\gamma=2}^{S \wedge \Gamma} \frac{S!(N-\gamma)!}{(S-\gamma)!N!}
$$
$$
+ \sum_{l=2}^{D} \frac{l(1-u)}{\Gamma-1} \left(\sum_{n=2}^{\Gamma} P(\Delta = l \,|\, \gamma = n)\right). \quad (11)
$$

D. Analysis of Throughput

In this section, we calculate the different throughput metrics and establish the stability condition for the network. There are N nodes in the network, each independently generating a new packet at the beginning of a frame with probability σ. Each generated packet requires on average $E[\Delta]$ packet copy transmissions. Thus, the network load in terms of packet copy transmissions per frame is $N \cdot \sigma \cdot E[\Delta]$ in the long-run average. Recalling that the AWG provides $D \cdot \Lambda$ wavelength channels, each providing one data phase per frame, we note that the network is stable if $N \cdot \sigma \cdot E[\Delta] < D \cdot \Lambda$.

For stable network operation (and negligible packet drop probabilities), the number of generated packets in a frame is equal to the number of completed packet transmissions (including all the required packet copy transmissions) in a frame in steady state. Hence, the multicast throughput is given by

$$
Z_M = N \cdot \sigma. \quad (12)
$$

Similarly, we obtain for the transmitter throughput in steady state

$$Z_T = N \cdot \sigma \cdot E[\Delta]. \tag{13}$$

The receiver throughput in steady state is given by

$$Z_R = N \cdot \sigma \cdot \left[u + (1 - u)\frac{\Gamma + 2}{2} \right] \tag{14}$$

because a given multicast packet with a maximum multicast size of Γ is received on average by $(\Gamma + 2)/2$ nodes.

E. Arrivals to Virtual Queue

In this section, we analyze the packet (copy) arrival to a given virtual queue representing a given AWG input–output port pair. That is, we study the arrivals to one (arbitrary) of the D virtual queues illustrated in Fig. 8.

There are $S = N/D$ nodes attached to the considered AWG input port. Each of the S nodes generates traffic mutually independent of the other nodes. Recall that a given node generates a new unicast data packet with probability $\sigma_u = \sigma \cdot u$ at the beginning of a given frame. With probability $1/D$, that packet is destined to the considered virtual queue.

Next, recall that a given node generates a new multicast packet with probability $\sigma_m = \sigma \cdot (1 - u)$ at the beginning of a frame. The number of destination nodes γ is uniformly distributed over $(2, \Gamma)$, and the individual destination nodes are uniformly distributed over the network nodes (and consequently AWG output ports and thus virtual queues). Given a multicast packet with γ destination nodes, we need to evaluate the probability that a packet copy is placed in the considered virtual queue. To evaluate this probability, we consider now a fixed virtual queue, say the queue associated AWG output port 1, or equivalently, color 1 in the urn model. The event that the multicast packet has at least one destination at AWG output port 1 corresponds to the events $C_{k_1,\ldots,k_D}$ with $0 \le k_1 \le \gamma \wedge S; 0 \le k_i \le \gamma \wedge S;$ for $i = 2, \ldots, D$, and $\sum_{i=1}^{D} k_i = \gamma$ in our urn model (1). Thus, the probability that a given multicast packet with γ destinations has at least one destination at the considered AWG output port is

$P(\text{“multicast pkt w. } \gamma \text{ dest. has copy to queue 1”})$

$$= P\left(C_{k_1,\ldots,k_D}(0 < k_1 \le \gamma \wedge S; 0 \le k_i \le \gamma \wedge S; \right.$$
$$\left. i = 2, \ldots, D; \sum_{i=1}^{D} k_i = \gamma) \right)$$

$$= 1 - P\left(C_{k_1,\ldots,k_D}(k_1 = 0; 0 \le k_i \le \gamma \wedge S; \right.$$
$$\left. i = 2, \ldots, D; \sum_{i=2}^{D} k_i = \gamma) \right)$$

$$= 1 - \sum_{\substack{0 \le k_2,\ldots,k_D \le \gamma \wedge S \\ \sum_{i=2}^{D} k_i = \gamma}} \frac{\binom{S}{k_2} \cdots \binom{S}{k_D}}{\binom{N}{\gamma}} \tag{15}$$

$$= 1 - \frac{(N - \gamma)!(N - S)!}{(N - \gamma - S)!N!}$$
$$\times \sum_{\substack{0 \le k_2,\ldots,k_D \le \gamma \wedge S \\ \sum_{i=2}^{D} k_i = \gamma}} \frac{\binom{S}{k_2} \cdots \binom{S}{k_D}}{\binom{N - S}{\gamma}} \tag{16}$$

$$= 1 - \frac{(N - \gamma)!(N - S)!}{(N - \gamma - S)!N!}. \tag{17}$$

Note that we obtained (17) by noting that the sum in (16) is over a complete set of events and by noting that for $\gamma > N - S$, there is a destination of the multicast at port 1 with probability one. Expressions (16) and (17) are valid for $\gamma \le N - S$. Now considering jointly the possibilities that a generated packet is a unicast packet or a multicast packet, the probability that a given node generates a packet (copy) for the considered queue in a given frame is

σ_q

$$= \frac{\sigma u}{D} + \frac{\sigma(1 - u)}{\Gamma - 1} \left[\sum_{\gamma = 2}^{\min\{\Gamma, N - S\}} \left(1 - \frac{(N - \gamma)!(N - S)!}{(N - \gamma - S)!N!} \right) \right.$$
$$\left. + \sum_{\gamma = N - S + 1}^{\Gamma} 1 \right] \tag{18}$$

$$= \frac{\sigma u}{D} + \frac{\sigma(1 - u)}{\Gamma - 1} \left[\sum_{\gamma = 2}^{\min\{\Gamma, N - S\}} \left(1 - \frac{(N - \gamma)!(N - S)!}{(N - \gamma - S)!N!} \right) \right.$$
$$\left. + (\Gamma - N + S)^+ \right] \tag{19}$$

where $(x)^+ = \max(0, x)$. Let A be a random variable denoting the number of packet (copy) arrivals to the considered virtual queue in a given frame. Let $a_i = P[A = i], i = 0, 1, \ldots, S$ denote the distribution of A. Clearly with S independent nodes generating traffic for the considered queue

$$a_i = \binom{S}{i} \cdot \sigma_q^i \cdot (1 - \sigma_q)^{(S - i)} \tag{20}$$

for $0 \le i \le S$ and $a_i = 0$ for $i > S$. We remark that the average number of packet copies generated by the S nodes attached to a given AWG input port in a frame equals the average number of packet copies arriving to the D virtual queues connecting the input port to the D AWG output ports in a frame, that is, $S \cdot \sigma \cdot E[\Delta] = S \cdot \sigma_q \cdot D$, which gives a convenient alternative expression for $E[\Delta]$.

F. Queuing Analysis of Virtual Queue

In this section, we conduct a queueing analysis of the virtual queue to determine the expected queue length and subsequently the different delay metrics. We begin our formulation by first noting that the arrival process is independent from the state of the queue. Second, we note that the arrival process in frame $t + 1$ denoted by A_{t+1} is independent of the arrival process A_t in the prior frame t. Let X_t denote the number of packet

(copies) in the queue at the beginning of a given frame t before the new packets are generated for the frame. We impose a maximum virtual queue occupancy J for calculation convenience and set it so large that boundary effects are negligible, that is, the occupancy J is not reached for stable operation. In each frame, up to R packets are served, that is, $X_{t+1} = \min[(X_t + A_t - R)^+, J]$. Thus, $(X_t)_{t \geq 0}$ is a Markov chain with state space $\mathcal{E} := \{0, 1, \ldots, J\}$ and the following transition matrix $\mathbf{P} = (p(x, y))_{x,y \in \mathcal{E}}$ with

$$p(x, 0) = \begin{cases} \sum_{i=0}^{R-x} a_i, & \text{for } x \leq R \\ 0, & \text{for } R < x \leq J \end{cases} \quad (21)$$

and

$$p(x, y) = \begin{cases} a_{R+y-x}, & \text{for } x \leq R + y \\ 0, & \text{for } x > R + y \end{cases} \quad (22)$$

for $0 < y \leq J - 1$ and

$$p(x, J) = P(A \geq R + J - x) = \sum_{i=R+J-x}^{N} a_i. \quad (23)$$

From (21)–(23), it follows that $\mathbf{P}$ is an aperiodic and irreducible transition matrix; hence, the Markov chain has an unique stationary probability distribution $\boldsymbol{\pi} = [\pi_0, \pi_1, \ldots, \pi_J]$ on $\mathcal{E}$ with $\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}$.

The expected queue length $E[X]$ is given by

$$E[X] = \sum_{j=1}^{J} j \cdot \pi_j. \quad (24)$$

We apply Little's theorem to find the mean copy delay

$$W_{\text{TR}} = \frac{E[X]}{S \cdot \sigma_q}. \quad (25)$$

To analyze the mean delay W_M, we need to consider the longest among the Δ virtual queues that a packet copy is placed in for a given generated packet. This analysis is complicated by the fact that multicasts with multiple packet copies destined to multiple queues in parallel tend to introduce correlations among the D virtual queues associated with a given AWG input port, whereby the larger the number of packet copies Δ, the stronger the correlation. If $\Delta = D$ with a high probability, then the D virtual queues behave essentially identically.

For the analytical evaluation of W_M, we need to note that the queueing model developed in this section considers a given virtual queue in isolation, that is, independently of the other $D - 1$ queues associated with the considered AWG input port. To evaluate W_M based on the developed queueing model, we employ the following heuristic. If Δ is below a threshold $\kappa \cdot D$ ($<D$), then we evaluate the longest queue with the order statistics of Δ independent virtual queues. If Δ is above the threshold $\kappa \cdot D$, then we approximate the longest queue by the length of one given independent virtual queue.

More formally, let $\hat{X}$ be a random variable denoting the number of packet copies in the longest queue that a given multicast feeds into in a given frame in steady state. Let $X_{[\delta]}$ be a random variable denoting the longest among $\Delta = \delta$

(independent) queues in steady state. From order statistics, we obtain that approximately

$$P(X_{[\delta]} = j) = \delta \cdot \left[\sum_{l=1}^{j} \pi_l \right]^{\delta-1} \cdot \pi_j. \quad (26)$$

Hence, approximately

$$E[\hat{X}] = \sum_{\delta=1}^{\kappa \cdot D} \left(\sum_{j=1}^{J} j \cdot P(X_{[\delta]} = j) \right) \cdot P(\Delta = \delta)$$
$$+ E[X] \cdot \sum_{\delta=\kappa \cdot D+1}^{D} P(\Delta = \delta) \quad (27)$$

where we assume that $\kappa \cdot D$ is an integer. Applying Little's theorem, we obtain the approximate mean multicast delay

$$W_M = \frac{E[\hat{X}]}{S \cdot \sigma_q}. \quad (28)$$

So far, we have assumed that the propagation delay in the network is negligible. We now outline how to incorporate propagation delay into our model. We assume that all nodes are equidistant from the central AWG (which can be achieved with fiber delay lines). We let τ denote the one-way end-to-end propagation delay in frames. We assume that the delay incurred for computing the data packet schedule is negligible (if significant, this delay could also be accounted for by τ). In the network with the TDMA control packet transmission and infinite scheduling window considered in this section, each packet incurs a delay of τ from its generation until the receipt of the corresponding control packet by all nodes and the successful scheduling of the data packet (copies). During this delay period, the data packet needs to be stored in the node (which we account for in the node buffer dimensioning in Section VI) and cannot yet be serviced. The data packet copies then incur the delay (W_{TR}), as calculated previously, from the time the transmission schedule has been computed until any arbitrary packet copy commences its transmission; analogously, the data packet incurs the delay W_M from the time the transmission schedule has been compared until the last copy of the packet commences. A given data packet copy incurs a transmission delay equal to the duration of the data phase (which we may roughly approximate by one frame) and a propagation delay τ for the propagation to the destination node. Thus, we need to add $2 \cdot \tau + 1$ frames to the queueing delays W_M and W_{TR} calculated above in order to account for the propagation delay.

V. THROUGHPUT-DELAY PERFORMANCE RESULTS

In this section, we give an overview of our numerical studies of the throughput-delay performance of the FT$^{\Lambda}$–FR$^{\Lambda}$ AWG network for unicast traffic, multicast traffic, as well as a mix of unicast and multicast traffic, whereby we primarily focus on illustrating the tradeoffs in selecting the AWG degree D and the number of utilized FSRs R for a given (fixed) number of transceivers $\Lambda = D \cdot R$, which is one of the key considerations in dimensioning the network. We refer the interested reader to [55] for more details. Initially, we fix the number of network nodes at

TABLE I
NETWORK PARAMETERS AND THEIR DEFAULT VALUES

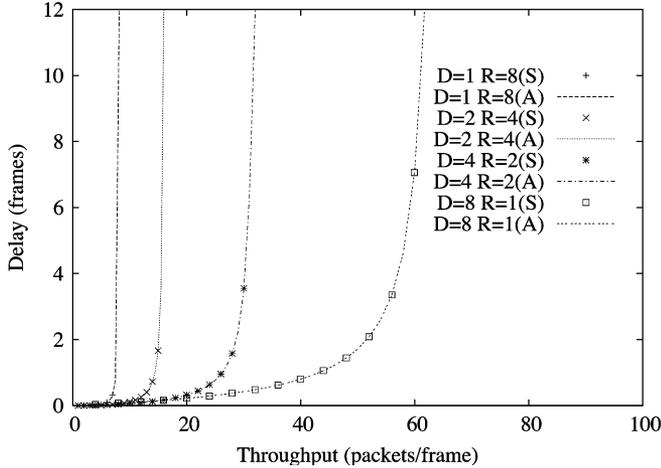| N	# of nodes in network	200
D	degree (# of ports) of AWG	1,2,4,8
R	number of utilized FSRs	1,2,4,8
Λ	$= D \cdot R$, # of wavelengths = # of transceivers in node	8
σ	packet generation probability	
u	fraction of unicast traffic	1, 0, 0.8
$1 - u$	fraction of multicast traffic	0, 1, 0.2
Γ	max # of dest. of multicast pkt	5, 15, 50, 200

Fig. 9. Delay W_M as a function of throughput Z_M for unicast traffic ($u = 1$).

Fig. 10. Delay W_M as a function of multicast throughput Z_M for multicast traffic $(1 - u = 1)$ with $\Gamma = 5$ and $\Gamma = 200$.

Fig. 11. Copy delay W_{TR} as a function of receiver throughput Z_R for multicast traffic ($u = 0$) with $\Gamma = 5$ and $\Gamma = 200$.

$N = 200$ and the number of used wavelengths (transceivers at each node) at $\Lambda = 8$. The network parameters are summarized in Table I. We assume that the propagation delay is negligible and that the node buffer and scheduling window are infinite at each node; finite buffers are studied in Section VI. We plot the numerical results from the probabilistic analysis (A), as well as simulation results (S). Each simulation was warmed up for 10^5 frames and terminated when the 99% confidence intervals of all performance metrics are less than 1% of the corresponding sample means.

A. Unicast Traffic

In Fig. 9, we plot the delay as a function of the throughput for different network configurations with $D \cdot R = \Lambda$ for unicast traffic ($u = 1$). In all these cases, the network has $\Lambda = 8$ wavelengths and $\Lambda = 8$ transceivers at each node. Note that the configuration ($D = 1, R = 8$) is equivalent to a PSC-based network. We observe that the ($D = 8, R = 1$) network has the largest throughput of up to 64 packets per frame. This better performance for larger D (with $\Lambda = D \cdot R$ fixed) for unicast traffic is due to spatial wavelength reuse, which results in a total of $D \cdot \Lambda$ wavelength channels between the AWG input and output ports.

B. Multicast Traffic

In Figs. 10 and 11, we plot the throughput and delay for multicast traffic ($u = 0$) for the ($D = 8, R = 1$) and ($D = 1, R = 8$) networks for different maximum multicast group sizes Γ. We
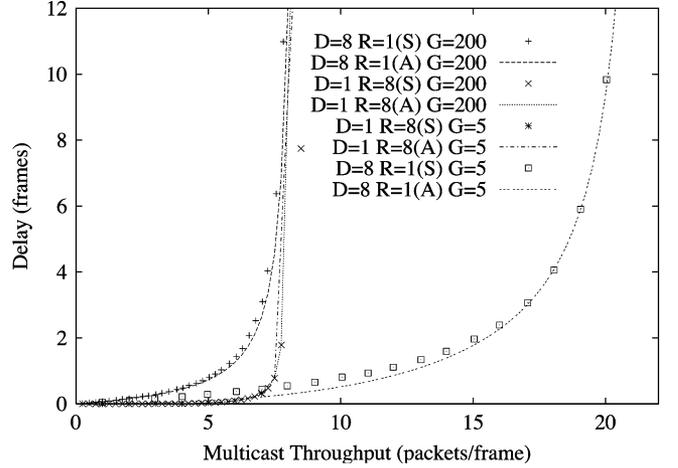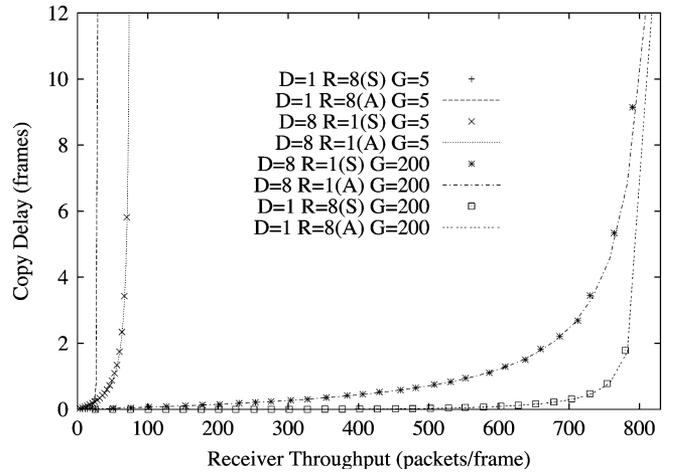
observe that as Γ increases, both network configurations converge to 1) a maximum multicast throughput of eight packets per frame and 2) the maximum receiver throughput of 800 packets per frame. To understand these dynamics, consider the transmission of broadcast packets that are destined to all $N = 200$ receivers in both networks. Clearly, in the PSC-equivalent ($D = 1, R = 8$) network, at most eight packet transmissions can take place simultaneously, each reaching all 200 receivers. In the ($D = 8, R = 1$) network, the broadcast of one packet requires the transmission of eight packet copies, one to each AWG output port, and reaching $N/D = 25$ receivers. Thus, in both networks, the multicast throughput, that is, the number of completed multicasts per frame, is eight packets per frame, and the receiver throughput is 1600 packets per frame. (Note that in this broadcast scenario, the transmitter throughput is eight packets per frame in the ($D = 1, R = 8$) network and 64 packets per frame in the ($D = 8, R = 1$) network.) Now with multicast traffic, with a maximum multicast group size of $\Gamma = 200$, a multicast packet has on average 100 destination nodes. The probability that at least one of these destination node is attached to each AWG output port is $P(\Delta = D \mid \gamma = 100) = 0.98$. Thus,
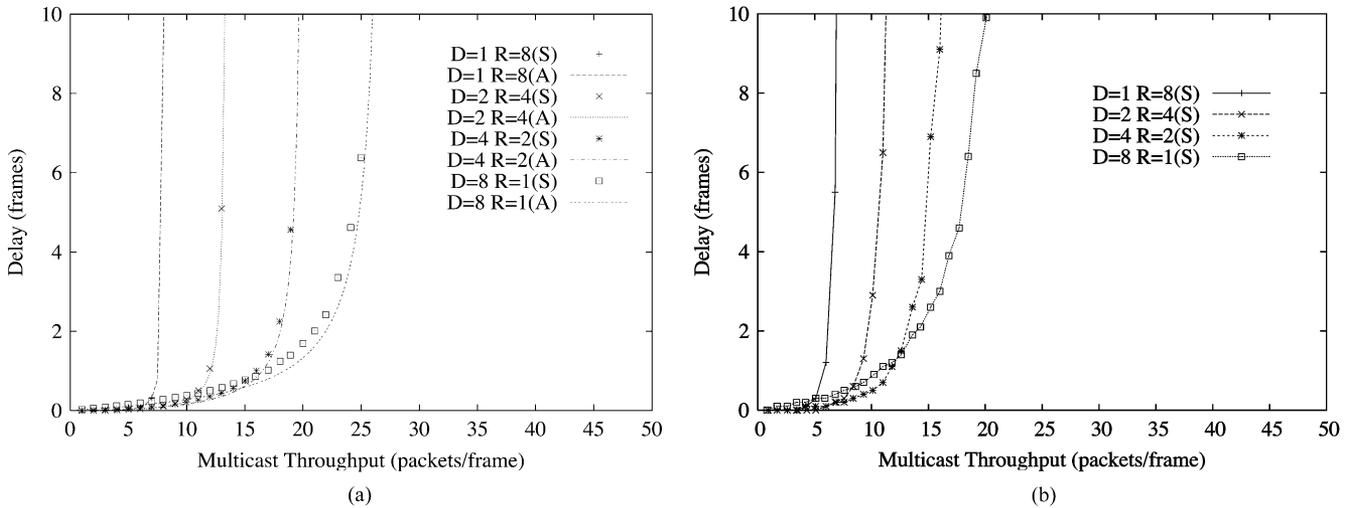
Fig. 12. Delay W_M as a function of multicast throughput Z_M for mix of 80% unicast ($u = 0.8$) and 20% multicast traffic with $\Gamma = 200$. (a) Bernoulli traffic. (b) Self-similar traffic.

it is very likely that D copies of the multicast packet need to be transmitted.

In general, when multicasting over the FT^Λ–FR^Λ AWG network, there are two effects at work. On one hand, a large AWG degree D increases the spatial wavelength reuse as all Λ wavelengths can be reused at each AWG port. On the other hand, as the multicast group size increases, it becomes (for uniformly distributed destination nodes) increasingly likely that at least one destination node is located at each AWG output port. The increase in spatial wavelength reuse in the network configuration with larger D is thus compensated by the increase in the number of required packet copy transmissions when the multicast group size is large. There is a net effect gain in the throughput performance whenever the number of required copy transmissions is smaller than the spatial reuse factor D, that is, when the multicast group size is relatively small or when the destination nodes tend to be co-located at a small number of AWG output ports. Indeed, as we see from Fig. 11, for a maximum multicast group size of $\Gamma = 5$ and a copy delay of two frames, the $(D = 8, R = 1)$ network achieves roughly twice the receiver throughput of the $(D = 1, R = 8)$ network.

Note that these multicast dynamics with transceiver arrays are fundamentally different from the dynamics with a single tunable transceiver at each node. In the single-transceiver network [6], [19], large multicasts are very difficult to schedule as it becomes increasingly unlikely to find the receivers of all destination nodes to be free at the same time, resulting in the so-called receiver bottleneck. Hence, it is advantageous to partition multicast groups into several smaller subgroups and transmit copies to each subgroup. The increased number of copy transmissions may lead to a channel bottleneck on the PSC, which can be relieved by the increased number of wavelength channels obtained from spatial wavelength reuse on the AWG. The increased number of transmissions on these larger number of channels in turn can exacerbate the receiver bottleneck with single-transceiver nodes [6], [19].

Returning to multicasting with transceivers arrays, which overcome the receiver bottleneck, we observe from Figs. 10

and 11 that the $(D = 8, R = 1)$ network gives larger delays than the $(D = 1, R = 8)$ network for large multicast group sizes. This is because the multiple packet copy transmissions required for large multicast group sizes in the $(D = 8, R = 1)$ network are more difficult to schedule than the single packet transmission in the $(D = 1, R = 8)$ network. In more detailed investigations [55], we found that for the $(D = 8, R = 1)$ network, the average copy delay W_{TR} is for low to moderate loads, typically 75%–80% of the corresponding delay W_M for completing the transmission of all packet copies.

In summary, our results indicate that the FT^Λ–FR^Λ AWG network has significantly improved throughput performance compared with an equivalent PSC network for small multicast groups or co-located multicast destinations. For large multicast groups with uniformly distributed destinations, the PSC network achieves smaller delays.

C. Mix of Unicast and Multicast Traffic

In this section, we consider mixes of unicast and multicast traffic, which are likely to arise in metropolitan area networks. Throughout this section, we fix the maximum multicast size at $\Gamma = 200$. In Fig. 12, we plot the throughput-delay performance of the FT^Λ–FR^Λ AWG network for 80% unicast traffic and 20% multicast traffic for different network configurations. For this traffic mix scenario, we consider both the Bernoulli traffic generation described in Section IV-A as well as self-similar traffic generation. In particular, we generate self-similar packet traffic with a Hurst parameter of 0.75, by aggregating ON–OFF processes with Pareto-distributed on-duration and geometrically distributed off-duration [54]. We observe that with increasing AWG degree D, the network achieves significantly larger multicast throughputs while the delay is increased only very slightly (at lower throughput levels). The throughput levels of the $(D = 8, R = 1)$ configuration are approximately three times larger than for the PSC-equivalent $(D = 1, R = 8)$ configuration.

This performance improvement is due to the increased spatial wavelength reuse with increased D, which is only, to a small degree, compensated for by the increased number of multicast

packet copy transmission for that typical mixed traffic scenario. In the PSC-based network $(D = 1)$, each packet transmission occupies one of the Λ wavelength channels irrespective of whether the packet is a unicast or a multicast packet. In the AWG-based network $(D \geq 2)$, each of the Λ wavelength channels can be reused at each AWG port, that is, D times, and additional copy transmissions are only required when the destination nodes of a given packet are attached to multiple AWG output ports. Thus, a larger AWG degree is overall beneficial when a significant portion of the traffic is unicast traffic.

We also observe from Fig. 12 that for self-similar traffic, the packet delays are somewhat larger compared with the delays for Bernoulli traffic. This is because with self-similar traffic generation, the packets arrive typically in bursts, which result in larger backlogs and longer queuing delays for the packets making up the tail end of a burst. (The impact of the self-similar traffic on the buffer requirements is studied in Section VI.) Nevertheless, the overall performance trends, that is, generally larger throughput and slightly increased delay at low throughput levels for larger D, are very similar both for Bernoulli and self-similar traffic. Hence, we focus on Bernoulli traffic for the remainder of this section.

In more extensive investigations [55], we have considered different fractions of unicast traffic and observed that the gap in performance between the PSC-based network $(D = 1, R = 8)$ and the AWG-based network with $D = 8$ widens as the fraction of unicast traffic increases. For 90% unicast traffic, the $(D = 8, R = 1)$ network achieves about four and a half times the throughput of the $(D = 1, R = 8)$ network, although the receiver throughput levels are reduced overall for the larger portion of unicast traffic.

We observe that the accuracy of our probabilistic analysis is quite good overall. The discrepancies between the analytical and simulation results for the delay W_M for larger D are primarily due to the heuristic approximation (27) of the occupancy distribution of the longest queue, for which we set $\kappa = 0.75$ throughout this paper.

In Tables II and III, we summarize the results of the network performance for the various AWG configurations for different traffic conditions. The data entries are extrapolated from our simulation results. In Table II, we fix the delay at four frames and record the maximum multicast throughput. In Table III, we fix the copy delay at four frames and record the maximum receiver throughput. We observe that both in terms of multicast throughput and receiver throughput, the $(D = 8, R = 1)$ network outperforms the networks with small D. Overall, our results indicate that the performance of the network improves as D becomes larger. This demonstrates the advantages of the spatial wavelength reuse of the AWG. The performance gap narrows for multicast-only traffic as the average number of destination nodes increases, and for the $u = 0, \Gamma = 200$ scenario, the $(D = 1, R = 8)$ network gives the largest throughputs. However, for the considered mixed unicast and multicast traffic scenarios, both the multicast throughput and the receiver throughput improve significantly as D increases. Both the multicast throughput and the receiver throughput for the $(D = 8, R = 1)$ configuration are over three times that of the $(D = 1, R = 8)$ PSC network.

TABLE II
MULTICAST THROUGHPUT Z_M (IN PACKETS/FRAME) FOR DELAY W_M OF FOUR FRAMES

(D, R)	$(u = 1)$	$(u = 0)$ $\Gamma = 5$	$(u = 0)$ $\Gamma = 15$	$(u = 0)$ $\Gamma = 200$	$(u = 0.8)$ $\Gamma = 200$
$(1, 8)$	7.9	7.9	7.9	7.9	7.9
$(2, 4)$	15.8	9.4	8.1	7.6	12.6
$(4, 2)$	31.2	12.3	8.4	7.4	19.7
$(8, 1)$	60.4	18.7	9.2	7.1	26.9

TABLE III
RECEIVER THROUGHPUT Z_R (IN PACKETS/FRAME) FOR COPY DELAY W_{TR} OF FOUR FRAMES

(D, R)	$(u = 1)$	$(u = 0)$ $\Gamma = 5$	$(u = 0)$ $\Gamma = 15$	$(u = 0)$ $\Gamma = 200$	$(u = 0.8)$ $\Gamma = 200$
$(1, 8)$	8	27	62	785	160
$(2, 4)$	16	30	66	762	245
$(4, 2)$	31	41	73	750	397
$(8, 1)$	60	64	97	730	490

D. Comparison Between TT–TR AWG Network and FT$^\Lambda$–FR$^\Lambda$ AWG Network

In this section, we compare the throughput-delay performance of the FT$^\Lambda$–FR$^\Lambda$ AWG network with the TT–TR AWG network employing one tunable transceiver at each node. Specifically, we consider 1) a TT–TR AWG network where the control packets are transmitted with a light-emitting diode (LED) (as in [5] and [6]) over the AWG and 2) a TT–TR–FT–FR AWG network where the control packets are transmitted over a PSC with a separate FT–FR at each node and the wavelengths on the AWG are available for data transmission all the time. TDMA control packet transmission is employed in all networks. We employ greedy data packet scheduling in the TT–TR AWG networks, which schedules a data packet for transmission to an AWG output port if at least one of the intended receivers at the port is free. This may result in multiple transmissions of a given multicast packet to a given AWG output port. This greedy policy is a reasonable benchmark for our comparisons as it tends to alleviate the receiver bottleneck at the expense of an increased burden on the transmitters, which as we demonstrate in the next section is a reasonable strategy.

For this investigation, we consider a control packet length of 2 B (which corresponds to the slot length) and a data packet length of 1500 B. The delay is given in slots, and the throughput is given in steady state, that is, normalized by the ratio of data phase to total frame length. In Fig. 13, we plot the throughput-delay performances of the two types of TT–TR AWG networks for different (D, R) combinations and compare with the $(D = 1, R = 8)$ FT$^\Lambda$ – FR$^\Lambda$ AWG network, which gives the worst throughput-delay performance of all (D, R) combinations for the FT$^\Lambda$–FR$^\Lambda$ AWG network (see Fig. 12).

We observe that for the considered typical traffic mix, all configurations of the TT–TR–FT–FR AWG network, which represents the best possible performance of a TT–TR AWG network in that all control is conducted in parallel over the PSC, have significantly lower performance than the worst performing FT$^\Lambda$–FR$^\Lambda$ AWG network configuration. The large delays for the TT–TR AWG network are due to the LED control packet transmission, which is conducted in cycles of length D frames [5], [6].
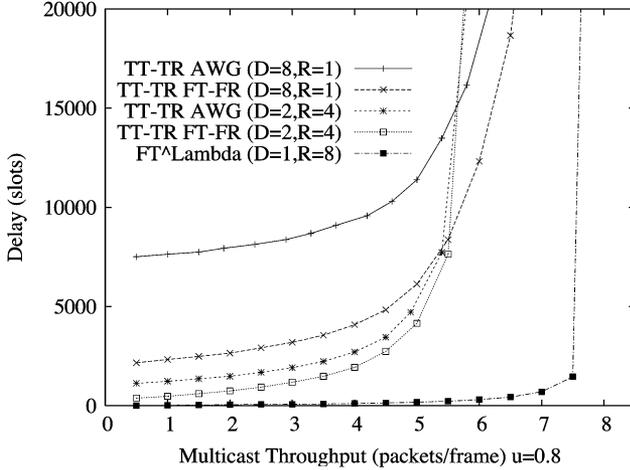
Fig. 13. Delay W_M as a function of multicast throughput Z_M for TT–TR AWG, TT–TR–FT–FR AWG, and FT$^\Lambda$–FR$^\Lambda$ AWG networks for a mix of 80% unicast ($u = 0.8$) and 20% multicast traffic with $\Gamma = 200$.

TABLE IV
TRANSCEIVER UTILIZATION COMPARISON FOR MIXED TRAFFIC ($u = 0.8$)
WITH $\Gamma = 200$ FOR DELAY OF 10 000 SLOTS

AWG Network	Z_T	U_T	Z_R	U_R
FT–FR–TT–TR ($D = 8, R = 1$)	22	0.11	126	0.63
FT^Λ ($D = \Lambda = 1$)	0.7	0.003	15	0.08
FT^Λ ($D = \Lambda = 2$)	3.6	0.01	60	0.15
FT^Λ ($D = \Lambda = 4$)	15	0.02	191	0.24
FT^Λ ($D = \Lambda = 8$)	61	0.04	535	0.33

E. Transceiver Utilization

In this section, we study the utilization of the transmitters and receivers in the FT$^\Lambda$–FR$^\Lambda$ and TT–TR AWG networks. We define the transmitter utilization U_T as the average fraction of time that any given transmitter is busy transmitting data packets in steady state. For the FT$^\Lambda$–FR$^\Lambda$ AWG network, clearly $U_T = Z_T/(N \cdot \Lambda) = \sigma \cdot E[\Delta]/\Lambda$. We define the receiver utilization U_R as the average fraction of time that any given receiver is busy receiving data packets in steady state. For the FT$^\Lambda$–FR$^\Lambda$ AWG network, clearly $U_R = Z_R/(N \cdot \Lambda) = \sigma \cdot [u + (1 - u) \cdot (\Gamma + 2)/2]/\Lambda$.

For the TT–TR AWG network, the transmitter utilization is difficult to compute because with the employed greedy scheduling algorithm, a packet copy destined to multiple receivers attached to the same splitter can be transmitted multiple times, depending on receiver availability. The receiver utilization for the TT–TR AWG network is approximately equal to the average number of destinations per packet multiplied by the packet throughput, that is, $U_R = \sigma \cdot [u + (1 - u) \cdot (\Gamma + 2)/2]$.

In Table IV, we compare the average transceiver utilization of the TT–TR–FT–FR AWG and the FT$^\Lambda$–FR$^\Lambda$ AWG networks for traffic loads resulting in an average delay of 10 000 slots. We observe that for the considered traffic mix with 80% unicast traffic and 20% multicast traffic, the utilization of the fixed-tuned transmitters in the FT$^\Lambda$–FR$^\Lambda$ AWG network is below 4% for all considered configurations. On the other hand, the fixed-tuned receivers are fairly well utilized, especially for the configurations with larger D. This suggests the need to

study TT^i–FR$^\Lambda$ AWG networks with $1 \le i < \Lambda$ in future work. This is further indicated by the utilization of approximately 11% of the tunable transmitter in the TT–TR–FT–FR AWG network. The tunable receiver in the TT–TR–FT–FR AWG network is heavily utilized, which illustrates the receiver bottleneck in TT–TR AWG networks and also indicates that an array of fixed-tuned receivers is a good choice for an AWG-based metro network carrying mixed traffic.

VI. NODE BUFFER DIMENSIONING

In this section, we address the problem of dimensioning the buffer in a node. Note that the analysis in Section IV considered virtual queues, whereby a virtual queue buffers the packet (copies) originating from the nodes attached to a given AWG input port and destined to nodes at a given AWG output port. We introduced the virtual queue as a modeling concept to make this analysis tractable. In a real network, the packets are buffered in node buffers. The dimensioning of these node buffers is important for network dimensioning and resource allocation. The probabilistic modeling of the nodal buffer occupancy is a complex problem due to the sharing of the wavelengths connecting a given AWG input–output port pair among the nodes connected to the input port and the multiple packet copies required to serve a multicast packet and is left for future work.

We conduct simulations of the FT$^\Lambda$–FR$^\Lambda$ network with the buffering at the nodes to determine the packet drop probability P_{loss}, which we define as the probability that a newly generated packet finds the nodal buffer full and is dropped. We denote L for the buffer capacity in the number of data packets at each node, whereby only one copy of each data packet is stored irrespective of the number of packet copy transmissions required to serve the packet. We consider the network with the default parameters given in Table I for a mix of 80% unicast traffic ($u = 0.8$) and 20% multicast traffic with $\Gamma = 200$. We consider both a network with a negligible propagation delay $\tau = 0$ and a network with a propagation delay of $\tau = 94$ frames, which corresponds to a typical scenario with a distance of 48.6 km between a node and the central AWG, a propagation speed of $2 \cdot 10^8$ m/s, a frame length of 1550 B, and an OC48 transmission rate of 2.4 Gb/s. In addition, we consider both Bernoulli (denoted ber) and self-similar (denoted ssim) traffic generation. In Fig. 14, we plot the packet drop probability P_{loss} at a node as a function of the probability σ that a node generates a new packet in a frame. We observe that for Bernoulli traffic and a negligible propagation delay, relatively small node buffers with a capacity of five or ten data packets are sufficient to achieve small loss probabilities on the order of 10^{-2} or less for traffic loads close to the stability limit of the networks in the considered scenario. Recall from Section IV-D that the stability limit for the network is $\sigma < D \cdot \Lambda/(N \cdot E[\Delta])$, which is $\sigma < 0.1$ for the considered $D = 4, R = 2$ network and $\sigma < 0.14$ for the considered $D = 8, R = 1$ network. For a propagation delay of $\tau = 94$ frames, correspondingly larger buffers are needed to store the data packets for which the control packets are propagating through the network. For self-similar traffic and a propagation delay of $\tau = 94$ frames, larger buffers are required to ensure small packet drop probabilities.
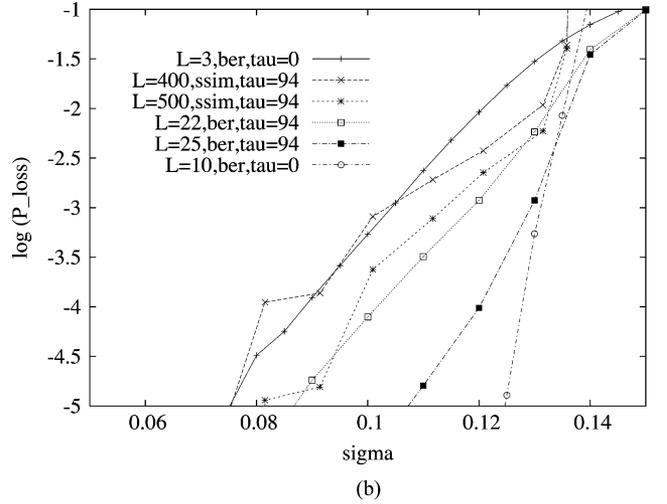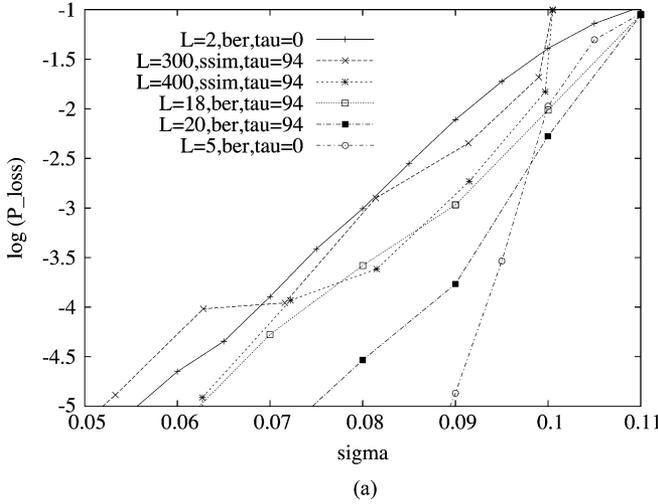
Fig. 14. Packet drop probability P_{loss} at node as a function of packet generation probability σ at node for different node buffer capacities L in packets. (a) $D = 4$ AWG ports and $R = 2$ FSRs. (b) $D = 8$ AWG ports and $R = 1$ FSR.

We observe from Fig. 14(b), however, that in the considered scenarios a buffer capable of holding 500 data packets ($=750$ kB for the considered 1500-B data packets) is sufficient to ensure loss probabilities below $10^{-3.5}$ for a long-run mean packet generation probability of $\sigma = 0.1$ (which for the considered network parameters corresponds to a long-run average traffic generation rate of 232 Mb/s of the bursty self-similar traffic with Hurst parameter $H = 0.75$).

TABLE V
PROBABILITY DISTRIBUTION AND EXPECTED VALUE OF NUMBER OF AWG OUTPUT PORTS WITH MULTICAST DESTINATIONS FOR $N = 20$ NODE NETWORK WITH $D = 4$ AND $S = 5$ FOR MULTICAST TRAFFIC ($u = 0.0$) WITH $\Gamma = 10$

	$P(\Delta = 1)$	$P(\Delta = 2)$	$P(\Delta = 3)$	$P(\Delta = 4)$	$E[\Delta]$
Urn with Repl.	0.037	0.222	0.371	0.371	3.075
Urn w/o Repl.	0.028	0.189	0.310	0.473	3.228
Simulation	0.027	0.190	0.309	0.473	3.228

VII. CONCLUSION

This paper described the development and evaluation of the FT$^\Lambda$–FR$^\Lambda$ AWG network, an AWG-based single-hop metro WDM network with a fixed-tuned transceiver-based node architecture. All building blocks of the network are well understood and commercially available, making the network practical and readily deployable. The analytical and simulation results in this paper indicate that the FT$^\Lambda$–FR$^\Lambda$ AWG network efficiently supports a typical mix of unicast and multicast traffic. For such a traffic mix, the FT$^\Lambda$–FR$^\Lambda$ AWG network with an 8×8 AWG achieves about three times the throughput of an equivalent PSC-based network.

There are several avenues for future work. One direction for future work is motivated by the finding in this paper that for a typical mix of unicast and multicast traffic, the utilization of the Λ fixed-tuned transmitters in the FT$^\Lambda$–FR$^\Lambda$ AWG network is relatively low, while the Λ fixed-tuned receivers are relatively highly utilized. This finding suggests the need to study AWG-based networks employing fast-tunable transmitters and arrays of fixed-tuned receivers. Such FT^i–FR$^\Lambda$ AWG networks with $1 \leq i < \Lambda$ also appear attractive from a technological perspective as fast-tunable transmitters are currently a relatively more mature technology compared with fast-tunable optical filter receivers. Note that the FT$^\Lambda$–FR$^\Lambda$ AWG network provides a useful benchmark for assessing the performance of AWG networks with other node architectures. This is because the data packet scheduling is only restricted by the wavelength channel resources in the FT$^\Lambda$–FR$^\Lambda$ AWG network. With other

node architectures, on the other hand, the data packet scheduling is typically limited by the channel as well as transceiver resources.

Another important direction for future work is to study efficient protection strategies for the AWG-based star networks.

APPENDIX A
EVALUATION OF $P(\Delta = l \mid \gamma = n)$

In this Appendix, we detail how to evaluate $P(\Delta = l \mid \gamma = n)$ given by (6). For $l = 2$, (6) takes the form

$$P(\Delta = 2 \mid \gamma = n) = \binom{D}{2} \sum_{\substack{1 \leq k_1, k_2 \leq n \wedge S \\ k_1 + k_2 = n}} \frac{\binom{S}{k_1}\binom{S}{k_2}}{\binom{N}{n}} \quad (29)$$

$$= \binom{D}{2} \sum_{k_1 = \max(1, n-S)}^{\min(n-1, S)} \frac{\binom{S}{k_1}\binom{S}{n-k_1}}{\binom{N}{n}}. \quad (30)$$

We define Q to represent the sum in (30), i.e.,

$$Q(\Delta = 2 \mid \gamma = n) = \sum_{k_1 = \max(1, n-S)}^{\min(n-1, S)} \binom{S}{k_1}\binom{S}{n-k_1}. \quad (31)$$

TABLE VI
PROBABILITY DISTRIBUTION AND EXPECTED VALUE FOR NUMBER OF AWG OUTPUT PORTS WITH MULTICAST DESTINATIONS FOR $N = 200$ NODE NETWORK WITH $D = 8$ AND $S = 25$ FOR MIX OF 80% UNICAST TRAFFIC ($u = 0.8$) AND 20% MULTICAST TRAFFIC WITH $\Gamma = 200$

	$P(\Delta = 1)$	$P(\Delta = 2)$	$P(\Delta = 3)$	$P(\Delta = 4)$	$P(\Delta = 5)$	$P(\Delta = 6)$	$P(\Delta = 7)$	$P(\Delta = 8)$	$E[\Delta]$
Replacement	0.800	0.001	0.002	0.002	0.003	0.004	0.008	0.180	2.351
Refined	0.800	0.001	0.002	0.002	0.003	0.004	0.007	0.181	2.353
Simulation	0.800	0.001	0.001	0.002	0.003	0.004	0.007	0.183	2.364

We note that when l increases by one in (6), we are adding one more term $\binom{S}{k_l}$. Thus

$$Q(\Delta = 3 \,|\, \gamma = n)$$
$$= \sum_{\substack{1 \le k_1, k_2, k_3 \le n \wedge S \\ k_1 + k_2 + k_3 = n}} \binom{S}{k_1}\binom{S}{k_2}\binom{S}{k_3} \tag{32}$$
$$= \sum_{k_3 = \max(1, n-2S)}^{\min(n-2, S)} Q(\Delta = 2 \,|\, \gamma = n - k_3) \binom{S}{k_3}. \tag{33}$$

In general

$$Q(\Delta = l \,|\, \gamma = n)$$
$$= \sum_{k_l = \max(1, n-(l-1)S)}^{\min(n-l+1, S)} Q(\Delta = l-1 \,|\, \gamma = n - k_l) \binom{S}{k_l}. \tag{34}$$

With the $Q(\Delta = l \,|\, \gamma = n)$, we can easily compute

$$P(\Delta = l \,|\, \gamma = n) = \binom{D}{l} \cdot \frac{Q(\Delta = l \,|\, \gamma = n)}{\binom{N}{n}}. \tag{35}$$

APPENDIX B
COMPARISON OF URN MODELS WITH AND WITHOUT REPLACEMENT FOR MULTICASTING

In this Appendix, we compare the urn model with replacement for the multicasting developed in [51] with the urn model without replacement developed in this paper. The urn model with replacement is simpler as it does not keep track of the balls that have already been drawn. Instead, when a ball (node) is drawn, the color (AWG output port) of the ball is noted, and the ball is put back into the urn. Then, the next ball is drawn, and so on. This urn model with replacement makes a modeling error in that it allows a given node to be drawn multiple times as a destination of a given multicast. In contrast, the urn model without replacement allows each node to be counted only once as a destination of a given multicast. To illustrate these effects, consider a network with $D = 2$ AWG input ports and $D = 2$ output ports, $N = 2$ nodes and $S = 1$ nodes attached to each AWG output port for multicast traffic ($u = 0$) destined to two nodes ($\gamma = 2$). Clearly, in this scenario, each packet is destined to both AWG output ports, that is, $P(\Delta = 2 \,|\, \gamma = 2) = 1$, as correctly modeled by the urn model without replacement. With the urn model with replacement, on the other hand, we obtain $P(\Delta = 1 \,|\, \gamma = 2) = 0.5$ and $P(\Delta = 2 \,|\, \gamma = 2) = 0.5$. To see

this, note that with probability 0.5, the ball selected in the second drawing is identical to the ball selected in the first drawing; with probability 0.5, the other ball is selected.

The modeling error of the urn model with replacement decreases as the probability of drawing the same ball multiple times decreases, which decreases as the number of balls (nodes in the network) increases. To illustrate the effect of the decreasing modeling error, we compare in Table V the probability distribution and expected value of the number of AWG output ports with attached destination nodes obtained from the urn model with replacement, the urn model without replacement, and simulations for a network with $N = 20$ nodes with $D = 4$ and $S = 5$ for multicast traffic $u = 0$ with a maximum of $\Gamma = 10$ destination nodes. We observe from the table that the urn model with replacement gives too-large values for the probabilities that the destinations are attached to a small number of AWG output ports and too-small values for the probability that the multicast destinations are attached to a large number of AWG output ports. The urn model with replacement thus gives too-small values overall for the expected number of AWG output ports with multicast destinations $E[\Delta]$. For the considered $N = 20$ node network, the urn model with replacement underestimates $E[\Delta]$ by almost 5%, which results in a correspondingly large underestimation of the transmitter throughput Z_T (13), the probability of generating a packet copy for a virtual queue σ_q (18), and the delays. We also observe from the table that the results obtained with the urn model with replacement closely match the simulation results.

In Table VI, we consider a $N = 200$ node network with $D = 8$ and $S = 25$ for a mix of 80% unicast traffic ($u = 0.8$) and 20% multicast traffic with $\Gamma = 200$. We observe from the table that the results from both urn models match the simulation results very closely. This is due to 1) the large fraction of unicast traffic for which the modeling error of selecting the same ball multiple times does not arise and 2) the large number of network nodes, which results in a small probability of selecting the same ball multiple times in the urn model with replacement.

ACKNOWLEDGMENT

REFERENCES

[1] B. Mukherjee, "WDM optical communication networks: Progress and challenges," *IEEE J. Select. Areas Commun.*, vol. 18, no. 10, pp. 1810–1824, Oct. 2000.

[2] J. Cai, A. Fumagalli, and I. Chlamtac, "The multitoken interarrival time (MTIT) access protocol for supporting variable size packets over WDM ring network," *IEEE J. Select. Areas Commun.*, vol. 18, no. 10, pp. 2094–2104, Oct. 2000.

[3] C. S. Jelger and J. M. H. Elmirghani, "Photonic packet WDM ring networks architecture and performance," *IEEE Commun. Mag.*, vol. 40, no. 11, pp. 110–115, Nov. 2002.

[4] A. Carena, V. D. Feo, J. Finochietto, R. Gaudino, F. Neri, C. Piglione, and P. Poggiolini, "RingO: A demonstrator of WDM optical packet network on a ring topology," *IEEE J. Select. Areas Commun. (Special Issue on Advances in Metropolitan Optical Networks)*, vol. 22, no. 8, pp. 1561–1571, Oct. 2004.

[5] M. Scheutzow, M. Maier, M. Reisslein, and A. Wolisz, "Wavelength reuse for efficient packet-switched transport in an AWG-based metro WDM network," *J. Lightw. Technol.*, vol. 21, no. 6, pp. 1435–1455, Jun. 2003.

[6] M. Maier, M. Scheutzow, and M. Reisslein, "The arrayed-waveguide grating based single-hop WDM network: An architecture for efficient multicasting," *IEEE J. Select. Areas Commun.*, vol. 21, no. 9, pp. 1414–1432, Nov. 2003.

[7] K. V. Shrikhande, I. M. White, M. Rogge, F.-T. An, A. Srivatsa, E. Hu, S.-H. Yam, and L. Kazovsky, "Performance demonstration of a fast-tunable transmitter and burst-mode packet receiver for HORNET," in *Proc. Optical Fiber Communication Conf. (OFC 2001)*, vol. 4, Anaheim, CA, Mar. 2001, pp. ThG-1–ThG-3.

[8] E. Chan, Q. N. Le, M. Beranek, Y. Huang, D. Koshinz, and H. Hager, "A 12-channel multimode fiber-optic 1.0625-Gb/s fiber channel receiver based on COTS devices and MCM-L/COB/BGA packaging," *IEEE Photon. Technol. Lett.*, vol. 12, no. 11, pp. 1549–1551, Nov. 2000.

[9] M. Ibsen, S. Alam, M. Zervas, A. Grudinin, and D. Payne, "8- and 16-channel all-fiber DFB laser WDM transmitters with integrated pump redundancy," *IEEE Photon. Technol. Lett.*, vol. 11, no. 9, pp. 1114–1116, Sep. 1999.

[10] M. Maier, *Metropolitan Area WDM Networks—An AWG Based Approach*. Norwell, MA: Kluwer, 2003.

[11] M. Maier, M. Reisslein, and A. Wolisz, "Toward efficient packet switching metro WDM networks," *Opt. Netw. Mag.*, vol. 3, no. 6, pp. 44–62, Nov./Dec. 2002.

[12] B. Mukherjee, "WDM-based local lightwave networks, Part I: Single-hop systems," *IEEE Network*, vol. 6, no. 3, pp. 12–27, May 1992.

[13] M. Bandai, S. Shiokawa, and I. Sasase, "Performance analysis of multicasting protocol in WDM-based single-hop lightwave networks," in *Proc. IEEE GLOBECOM*, Phoenix, AZ, Nov. 1997, pp. 561–565.

[14] A. Bianco, G. Galante, E. Leonardi, F. Neri, and A. Nucci, "Scheduling algorithms for multicast traffic in TDM/WDM networks with arbitrary tuning latencies," in *Proc. IEEE GLOBECOM*, San Antonio, TX, Nov. 2001, pp. 1551–1556.

[15] M. S. Borella and B. Mukherjee, "Limits of multicasting in a packet-switched WDM single-hop local lightwave network," *J. High Speed Netw.*, vol. 4, pp. 155–167, 1995.

[16] J. P. Jue and B. Mukherjee, "The advantages of partitioning multicast transmissions in a single-hop optical WDM network," presented at the IEEE Int. Conf. on Communication (ICC), Montreal, QC, Canada, Jun. 1997.

[17] T. Kitamura, M. Iizuka, and M. Sakuta, "A new partition scheduling algorithm by prioritizing the transmission of multicast packets with less destination address overlap in WDM single-hop networks," in *Proc. IEEE GLOBECOM*, vol. 3, San Antonio, TX, Nov. 25–29, 2001, pp. 1469–1473.

[18] H.-C. Lin and P. S. Liu, "A reservation-based multicast scheduling algorithm with reservation window for single-hop WDM network," in *IEEE Int. Conf. Networks*, Singapore, Sep. 2000, p. 493.

[19] H.-C. Lin and C.-H. Wang, "A hybrid multicast scheduling algorithm for single-hop WDM networks," *J. Lightw. Technol.*, vol. 19, no. 11, pp. 1654–1664, Nov. 2001.

[20] T.-L. Liu, C.-F. Hsu, and N.-F. Huang, "Multicast QoS traffic scheduling with arbitrary tuning latencies in single-hop WDM networks," in *Proc. IEEE Int. Conf. on Communications (ICC)*, New York, Apr. 2002, pp. 2886–2890.

[21] E. Modiano, "Random algorithms for scheduling multicast traffic in WDM broadcast-and-select networks," *IEEE/ACM Trans. Netw.*, vol. 7, no. 3, pp. 425–434, Jun. 1999.

[22] Z. Ortiz, G. N. Rouskas, and H. G. Perros, "Maximizing multicast throughput in WDM networks with tuning latencies using the virtual receiver concept," *Eur. Trans. Telecommun.*, vol. 11, no. 1, pp. 63–72, Jan./Feb. 2000.

[23] G. N. Rouskas and M. H. Ammar, "Multidestination communication over tunable-receiver single-hop WDM networks," *IEEE J. Select. Areas Commun.*, vol. 15, no. 3, pp. 501–511, Apr. 1997.

[24] L. Sahasrabuddhe and B. Mukherjee, "Probability distribution of the receiver busy time in a multicasting local lightwave network," in *Proc. IEEE Int. Conf. on Communications (ICC)*, Montreal, QC, Canada, Jun. 1997.

[25] S.-T. Sheu and C.-P. Huang, "An efficient multicasting protocol for WDM star-coupler networks," in *Proc. IEEE Int. Symp. Computers Communications*, Alexandria, Egypt, Jul. 1997, pp. 579–583.

[26] W.-Y. Tseng, C.-C. Sue, and S.-Y. Kuo, "Performance analysis for unicast and multicast traffic in broadcast-and-select WDM networks," in *Proc. IEEE Int. Symp. Computers Communications*, Red Sea, Egypt, Jul. 1999, pp. 72–78.

[27] A. Ding and G.-S. Poo, "A survey of optical multicast over WDM networks," *Comput. Commun.*, vol. 26, no. 2, pp. 193–200, Feb. 2003.

[28] A. Hamad and A. Kamal, "A survey of multicasting protocols for broadcast-and-select single-hop networks," *IEEE Network*, vol. 16, no. 4, pp. 36–48, Jul./Aug. 2002.

[29] J. He, S.-H. G. Chan, and D. H. K. Tsang, "Multicasting in WDM networks," *IEEE Communications Surveys Tutorials*, vol. 4, no. 1, 3rd Quarter 2002.

[30] M. McKinnon, G. Rouskas, and H. Perros, "Performance analysis of a photonic single-hop ATM switch architecture, with tunable trasmitters and fixed frequency receivers," *Perform. Eval.*, vol. 33, no. 2, pp. 113–136, Jul. 1998.

[31] L. Wang and K. Lee, "A WDM based virtual bus for universal communication and computing systems," in *Proc. IEEE Int. Conf. on Communication (ICC)*, Jun. 1992, pp. 888–894.

[32] K. Kato, A. Okada, Y. Sakai, K. Noguchi, T. Sakamoto, A. Takahara, A. Kaneko, S. Suzuki, and M. Matsuoka, "10-Tbps full-mesh WDM network based on a cyclic-frequency arrayed-waveguide grating router," in *Proc. Eur. Conf. Optical Communication (ECOC 2000)*, vol. 1, Munich, Germany, Sep. 2000, pp. 105–107.

[33] A. Okada, T. Sakamoto, Y. Sakai, K. Noguchi, and M. Matsuoka, "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 *Optical Fiber Communication Conf. (OFC 2001) Tech. Dig.*, Anaheim, CA, Mar. 2001, Paper ThG5.

[34] S. B. Alexander, R. Bondurant, D. Byrne, V. Chan, S. Finn, R. Gallager, B. Glance, H. Haus, P. Humblet, R. Jain, I. Kaminow, M. Karol, R. Kennedy, A. Kirby, H. Le, A. Saleh, B. A. Schofield, J. Shapiro, N. Shankaranarayanan, R. Thomas, R. Williamson, and R. Wilson, "A pre-competitive consortium on wide-band all-optical networks," *J. Lightw. Technol.*, vol. 11, no. 5, pp. 714–735, May/Jun. 1993.

[35] D. Banerjee, J. Frank, and B. Mukherjee, "Passive optical network architecture based on waveguide grating routers," *IEEE J. Select. Areas Commun.*, vol. 16, no. 7, pp. 1040–1050, Sep. 1998.

[36] K. Bengi, *Optical Packet Access Protocols for WDM Networks*. Norwell, MA: Kluwer, 2002.

[37] M. Chia, D. Hunter, I. Andonovic, P. Ball, I. Wright, S. Ferguson, K. Guild, and M. O'Mahony, "Packet loss and delay performance of feedback and feed-forward arrayed-waveguide gratings-based optical packet switches wtih WDM inputs-outputs," *J. Lightw. Technol.*, vol. 19, no. 9, pp. 1241–1254, Sep. 2001.

[38] B. Glance, I. P. Kaminow, and R. W. Wilson, "Applications of the integrated waveguide grating router," *J. Lightw. Technol.*, vol. 12, no. 6, pp. 957–962, Jun. 1994.

[39] A. Hill, S. Carter, J. Armitage, M. Shabeer, R. Harmon, and P. Rose, "A scalable and switchless optical network structure, employing a single 32×32 free-space grating multiplexer," *IEEE Photon. Technol. Lett.*, vol. 8, no. 4, pp. 569–571, Apr. 1996.

[40] D. K. Hunter and I. Andonovic, "Approaches to optical internet packet switching," *IEEE Commun. Mag.*, vol. 38, no. 9, pp. 116–122, Sep. 2000.

[41] D. K. Hunter, H. M. Nizam, M. C. Chia, I. Andonovic, K. Guild, A. Tzanakaki, M. O'Mahony, L. Bainbridge, M. Stephens, R. Penty, and I. White, "WASPNET: A wavelength switched packet network," *IEEE Commun. Mag.*, vol. 37, no. 3, pp. 120–129, Mar. 1999.

[42] D. Jung, S. Shin, C. Lee, and Y. Chung, "Wavelength-division-multiplexed passive optical network based on spectrum-slicing techniques," *IEEE Photon. Technol. Lett.*, vol. 10, no. 9, pp. 1334–1336, Sep. 1998.

[43] M. J. Karol and B. Glance, "A collision-avoidance WDM optical star network," *Comput. Netw. ISDN Syst.*, vol. 26, pp. 931–943, Mar. 1994.

[44] 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, no. 3, pp. 128–135, Mar. 2001.

[45] M. J. Spencer and M. Summerfield, "WRAP: A medium access control protocol for wavelength-routed passive optical networks," *J. Lightw. Technol.*, vol. 18, no. 12, pp. 1657–1676, Dec. 2000.

[46] A. Bianco, E. Leonardi, M. Mellia, and F. Neri, "Network controller design for SONATA—A large-scale all-optical passive network," *IEEE J. Select. Areas Commun.*, vol. 18, no. 10, pp. 2017–2028, Oct. 2000.

[47] 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," *Eur. Trans. Telecommun.*, vol. 11, no. 1, pp. 73–82, Jan./Feb. 2000.

[48] C. Fan, M. Maier, and M. Reisslein, "The AWG‖PSC network: A performance enhanced single-hop WDM network with heterogeneous protection," in *Proc. IEEE INFOCOM*, San Francisco, CA, Mar. 2003, pp. 2279–2289.

[49] A. Hill, M. Brierley, R. Percival, R. Wyatt, D. Pitcher, K. I. Pati, I. Hall, and J.-P. Laude, "Multi-star wavelength-router network and its protection strategy," *IEEE J. Select. Areas Commun.*, vol. 16, no. 7, pp. 1134–1145, Sep. 1998.

[50] Y. Sakai, K. Noguchi, R. Yoshimura, T. Sakamoto, A. Okada, and M. Matsuoka, "Management system for full-mesh WDM AWG-STAR network," in *Proc. Eur. Conf. Optical Communications (ECOC 2001)*, vol. 3, Amsterdam, The Netherlands, Sep. 2001, pp. 264–265.

[51] C. Fan, M. Reisslein, and S. Adams, "The $FT^\Lambda - FR^\Lambda$ AWG network: A practical single-hop metro WDM network for efficient uni- and multicasting," in *Proc. IEEE INFOCOM*, vol. 1, Mar. 2004, pp. 502–513.

[52] L. Takacs, *Combinatorial Methods in the Theory of Stochastic Processes*. New York: Wiley, 1967.

[53] N. L. Johnson and S. Kotz, *Urn Models and Their Applications*. New York: Wiley, 1977.

[54] K. Park and W. Willinger, Eds., *Self-Similar Network Traffic and Performance Evaluation*. New York: Wiley, 2000.

[55] C. Fan, "Architecture and MAC protocols for AWG-based WDM single-hop networks," Ph.D. dissertation, Dept. of Electrical Engineeering, Arizona State University, Tempe, 2004.
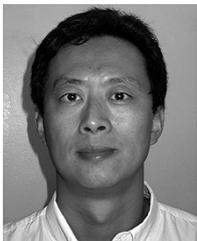
Chun Fan was born on August 14, 1968 in Beijing, China. He received the B.S. degree in industrial engineering from the University of Arizona, Tucson, in 1990 and the M.B.A. degree from American Graduate School of International Management, Thunderbird, AZ. He is currently working toward the Ph.D. degree in the Department of Electrical Engineering at the Fulton School of Engineering, Arizona State University, Tempe.

Prior to working toward the Ph.D. degree, he worked for Advanced Fiber Communications, Petaluma, CA, from 1998 to 2000 and for Emerson Electric from 1992 to 1997. His current research focuses on the design and performance evaluation of wavelength-division-multiplexing single-hop networks.

Stefan Adams received the Dipl.-Phys. degree from the University of Göttingen, Göttingen, Germany, in 1995 and the Dipl.-Math. degree from the University of Hagen, Hagen, Germany, in 1998, both in theoretical physics and mathematics. He received the Ph.D. in mathematics from the University of Munich, Munich, Germany, in 2000.

He is currently a Postdoctoral Researcher (Habilitant) in the Department of Mathematics at the Technical University Berlin, Berlin, Germany.

Dr. Adams is currently a Member of the Deutsche Forschungsgemeinschaft Research Center MATHEON "Mathematics for Key Technologies, " Berlin, Germany.

Martin Reisslein (A'96–S'97–M'98–SM'03) 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 from the University of Pennsylvania in 1998.

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 teaching courses on performance evaluation and computer networking at the Technical University Berlin, Berlin, Germany. He is currently an Assistant Professor in the Department of Electrical Engineering at Arizona State University, Tempe. 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://trace.eas.asu.edu. His research interests are in the areas of Internet Quality of Service, video traffic characterization, wireless networking, and optical networking.

Dr. Reisslein is co-recipient of the Best Paper Award of the SPIE Photonics East 2000—Terabit Optical Networking Conference. He is Editor-in-Chief of the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS and has served on the Technical Program Committees of IEEE INFOCOM, IEEE GLOBECOM, and the IEEE International Symposium on Computer and Communications. In addition, he has organized sessions at the IEEE Computer Communications Workshop (CCW).