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

Chun Fan
and Martin Reisslein
Dept. of Electrical Eng.
Arizona State Univerity
Tempe, AZ 85287–5706
Email: {chun.fan, reisslein}@asu.edu

Stefan Adams
Dept. of Mathematics
Technical University Berlin
10587 Berlin, Germany
Email: adams@math.tu-berlin.de

***Abstract*— Single–hop 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. In this paper we develop and evaluate 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. Our results from probabilistic analysis and simulation indicate that the $FT^{\Lambda} - FR^{\Lambda}$ AWG network gives particularly good throughput–delay performance for multicast traffic with small multicast group sizes or localized destination nodes, as well as for a mix of unicast and multicast traffic.**

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 WDM networks, the metropolitan area networks are becoming a bottleneck—the so called metro-gap. This is largely due to the current circuit–switched SONET/SDH over WDM metro networks, which carry the 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 (multi–destination) 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 tele–medicine. 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 multi–hop 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). 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 which 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 to PSC based networks [2].

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 low power consumption and small foot print, fast–tunable transceivers are generally a less mature technology than fixed–tuned transceiver arrays. More specifically fast–tunable transmitters have just recently been experimentally proven to be feasible in a cost–effective manner [3], 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 [4], [5], more mature, more reliable, and commercially available, but also have some drawbacks such as increased power consumption and larger footprint. At the MAC protocol level, transceiver arrays have a number of distinct advantages. The transmitter arrays allow for high–speed signaling over the AWG

in contrast to the low-speed signaling through the spectral slicing of broadband light sources [2] which suffer from a small bandwidth–distance product. The receiver arrays, on the other hand, relieve the receiver bottleneck caused by multicast traffic, that is transmitted over the large number of wavelength channels obtained from spatial wavelength reuse on the AWG.

In this paper we develop and evaluate the $FT^{\Lambda}-FR^{\Lambda}$ AWG network, an AWG based single–hop WDM network with an array of fixed–tuned transmitters and receivers at each network node. The proposed $FT^{\Lambda} - FR^{\Lambda}$ 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 $FT^{\Lambda} - FR^{\Lambda}$ 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. In the following subsection we review related work. In Section II we describe the architecture of the $FT^{\Lambda} - FR^{\Lambda}$ AWG network and discuss how it supports unicast and multicast traffic. In Section III we provide the distributed medium–access–control (MAC) protocol. In Section IV we develop a probabilistic model to evaluate the throughput–delay performance of the network for a mix of unicast and multicast traffic. In Section V we present numerical performance results obtained from our analytical model and simulation. We summarize our findings in Section VI.

A. Related Work

Both unicasting (see surveys [6], [7]) and multicasting (see surveys [8], [9]) over Passive Star Coupler (PSC) based networks have been studied extensively. The studied PSC based networks include networks with arrays of fixed–tuned receivers (see for instance [10]), as well as networks with arrays of fixed–tuned transmitters and receivers (see for instance [11]). 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 [12], [13]. General design principles for networks based on AWGs are studied, for instance, in [14].

SONATA [15], [16] 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, we consider a metropolitan area network in this paper with decentralized medium access control. Our 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 [2], [17]. The network considered in [2], [17] employs a single fast–tunable transmitter and a single fast–tunable receiver at each node.

We remark that we focus on the network and MAC protocol design of the $FT^{\Lambda}-FR^{\Lambda}$ AWG network and its performance evaluation in this paper. The protection and survivability aspects of the network are beyond the scope of this paper. We note that protection strategies for AWG based networks have been examined in [18], [19], [20]. In our ongoing work we are developing similar strategies for the $FT^{\Lambda} - FR^{\Lambda}$ network.

II. Architecture

Our AWG based network architecture is illustrated in Fig. 1. The AWG has D input ports and D output ports. 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 $N_{i,j}$.

The wavelength routing property of the AWG is illustrated for a 2×2 AWG with a period of the wavelength response (referred to as Free Spectral Range (FSR)) of $R = 2$ in Fig 2. According to the periodic wavelength routing, every second 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

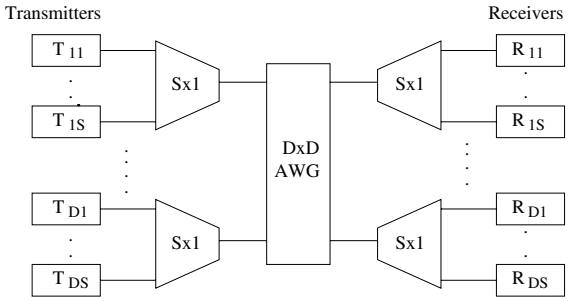

Fig. 1. Network architecture

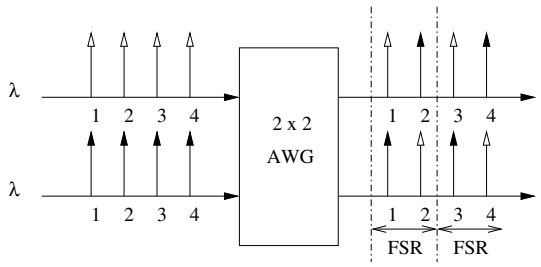

Fig. 2. Wavelength routing in 2×2 AWG with $R = 2$ FSRs

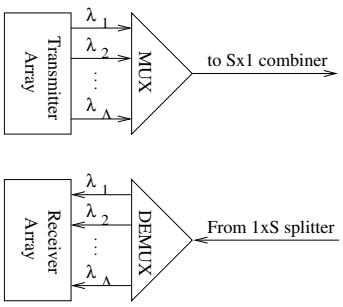

Fig. 3. Detailed node architecture

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.

We close 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, i.e., a packet that is destined to one destination node, requires one transmission on the wavelength that is routed to the AWG output port that the destination node is attached to.

Now consider a multicast packet, i.e., 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 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. 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

The $FT^{\Lambda} - FR^{\Lambda}$ AWG network uses pre–transmission coordination together with global scheduling to coordinate the access of the nodes to the shared wavelength channels. This coordination and scheduling are generally recommended strategies for achieving good throughput delay performance in shared-wavelength single-hop star networks [6]. Time is divided into frames; each frame consists 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 outline 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 $d = 1, 2, \ldots, D$, transmits its control packet. Each node uses its full array of fixed transmitters for high–speed control packet transmission (in contrast to the lower speed signaling with spreading and spectral slicing employed in the single transceiver network [2]). 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. 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 $S = N/\Lambda$.)

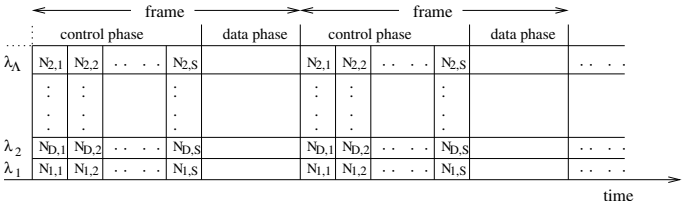

Fig. 4. Frame structure and control packet reception schedule for nodes at AWG output port 1 of network with $R = 1$ FSR

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. 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

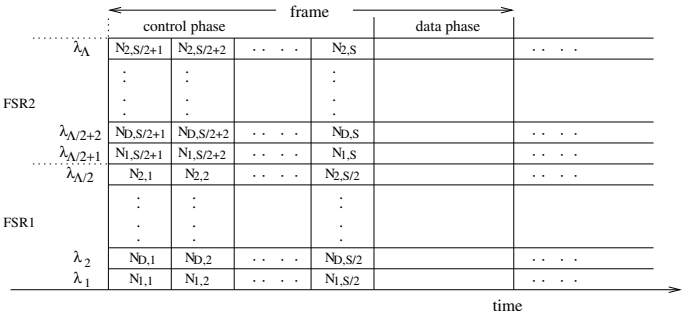

Fig. 5. Frame structure and control packet reception schedule for nodes at AWG output port 1 of network with $R = 2$ FSRs

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

In a network with $R = 1$ FSR, each node sends the control packet uniformly and randomly in of the M, $M \leq N/\Lambda$, control slots 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 and uniformly to transmit the control packet, as illustrated in Fig. 5 for R = 2.

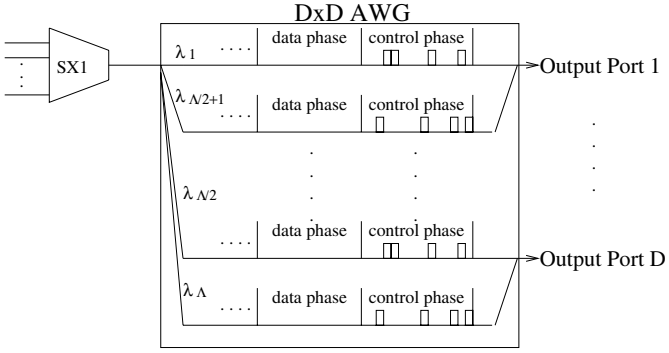

Fig. 6. Control packet contention and frame structure for network with $R = 2$ FSRs

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.

We note that in the $FT^{\Lambda} - FR^{\Lambda}$ 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 to 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 acknowledgements would be required to verify whether a control packet collision occurred; these acknowledgements would result in increased protocol complexity and delay.

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. For each unicast and multicast packet (copy) transmission, a wavelength is assigned on a first–come–first–served (FC–FS) basis starting with the lowest FSR in the immediate frame. We adopt the FC–FS scheduling since scheduling algorithms for high-speed WDM networks need to be of low complexity [21]. If the FSRs of the immediate frame are scheduled, then slots on the subsequent frame are assigned, and so on, up to a pre-specified scheduling window. If the data packet 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. To avoid unfairness, which may arise with the FC–FS scheduling if the control packets are transmitted in the fixed TDMA sequence and the scheduling window is small, the received control packets can be randomly resequenced before the scheduling commences.

Note that the data packets are buffered in the electronic domain at each source node which can have quite large memory capacity. We reasonably assume that the nodal buffer is large enough to hold all scheduled packets. Unscheduled packets may be dropped and are indicative of congestion. We leave traffic congestion management to the upper layer protocols.

IV. Analysis

A. Overview of System 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. 7. 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.

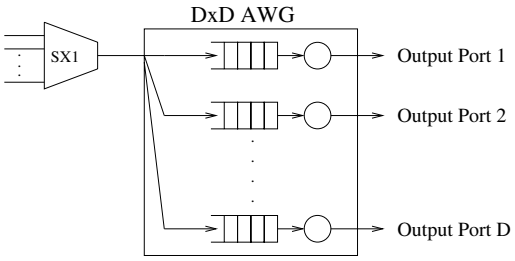

Fig. 7. Queuing model: one *virtual* queue for each AWG input–output port pair. Note that there is no physical buffer at the AWG.

We consider the following scenario in our modelling of the $FT^{\Lambda} - FR^{\Lambda}$ AWG network.

- *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 Γ, $2 \leq \Gamma \leq N$, 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 $2 \leq \gamma \leq \Gamma$, which is uniformly distributed, i.e., $\gamma \sim U(2, \Gamma)$.
- *Propagation delay*: The propagation delay is assumed to be negligible.
- *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*
- *Scheduling window*: For our analytical study of the throughput and delay metrics, as defined in the following subsection, we consider an infinite scheduling window and infinite nodal buffers, such that no packet is dropped. However, we study the network for stable operations, as detailed in Section IV-D.

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 modelling these D input–output port pairs.

B. Definition of Performance Metrics

In our 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: (i) the end of the control phase of the frame in which a packet is generated, and (ii) 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 $W_M = W_{TR}$. Also note that all of these performance 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. 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. 7.

There are $S = N/D$ nodes attached to the considered AWG input port. Each of the S nodes generates traffic mutually independently 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, the probability that a packet copy is placed in the considered virtual queue is $1-(1-1/D)^\gamma$. To see this, note that a given destination node (any of the γ nodes) is not attached to the considered queue (output port) with probability $1-1/D$. Hence, none of the γ destination nodes is attached to the considered queue with probability $(1-1/D)^\gamma$. The complementary event, that at least one of the γ destinations is attached to the considered queue, has thus the probability $1-(1-1/D)^\gamma$. Noting that γ is uniformly distributed over $(2, \Gamma)$, we obtain $\frac{1}{\Gamma-1}\sum_{\gamma=2}^{\Gamma}\left[1-\left(1-\frac{1}{D}\right)^\gamma\right]$ as the probability that a given multicast packet with a maximum number of Γ destinations results in a packet copy for the considered virtual queue. Note that the preceding model is approximate in that it considers the destinations of a given multicast as independent. A more elaborate model with accounts for the dependencies between the destinations of a given multicast is developed in [22].

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

$$\sigma_q = \sigma \cdot \left(\frac{u}{D} + \frac{1-u}{\Gamma-1} \sum_{\gamma=2}^{\Gamma} \left[1 - \left(1 - \frac{1}{D} \right)^\gamma \right] \right). \tag{1}$$

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,

$$a_i = \binom{S}{i} \cdot \sigma_q^i \cdot (1-\sigma_q)^{(S-i)}, \tag{2}$$

for $0 \le i \le S$ and $a_i = 0$ for $i > S$.

D. Analysis of Throughput

In this section we calculate the different throughput metrics and establish the stability condition for the network. First, we evaluate the number of packet copy transmissions required to service a given generated packet. Let Δ, $1 \le \Delta \le D$, 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 (i) the generated packet is a unicast packet (which has probability u), or (ii) the generated packet is a multicast packet (which has probability $1-u$) and all γ receivers are attached to the same AWG output port. Hence, for $\gamma \le S$,

$$P(\Delta = 1|\gamma) = u + (1-u) \cdot \frac{1}{D^{\gamma-1}}. \tag{3}$$

To see this, note that all γ destination nodes of a given multicast packet are attached to a given virtual queue with probability $1/D^\gamma$, thus with D/D^γ all destinations are attached to any of the D ports.

For $P(\Delta = \delta)$, $\delta = 2, \ldots, D$, we obtain

$$P(\Delta = \delta|\gamma) = (1-u) \cdot \binom{D}{D-\delta} \cdot \left(1 - \frac{D-\delta}{D}\right)^\gamma \cdot \sum_{l=0}^{\delta-1} \binom{\delta}{l} (-1)^l \left(1 - \frac{l}{\delta}\right)^\gamma \tag{4}$$

by noting that the distribution of the number of required packet copies is equivalent to the number of urns containing at least one ball when throwing δ balls uniformly randomly into D urns [23]. The expected number of required packet copy transmissions is then

$$E[\Delta] = \sum_{\delta=1}^{D} \delta \cdot \left[\frac{1}{\Gamma-1} \sum_{\gamma=2}^{\Gamma} P(\Delta = \delta|\gamma) \right]. \tag{5}$$

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, 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

$$Z_M = N \cdot \sigma. \tag{6}$$

Similarly, we obtain for the transmitter throughput in steady state

$$Z_T = N \cdot \sigma \cdot E[\Delta]. \tag{7}$$

The receiver throughput in steady state is given by

$$Z_R = N \cdot \sigma \cdot \left[u + (1-u)\frac{\Gamma+2}{2} \right], \tag{8}$$

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

E. Queuing Analysis of Virtual Queue

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. We measure the queue length immediately after the

TABLE I

SYSTEM OF EQUATIONS FOR STATIONARY PROBABILITIES

$\pi_0 =$	$\sum_{i=0}^{R} a_i \pi_0 + \sum_{i=0}^{R-1} a_i \pi_1 + \cdots + a_0 \pi_R$
$\pi_1 =$	$a_{R+1} \pi_0 + a_R \pi_1 + \cdots + a_1 \pi_R + a_0 \pi_{R+1}$
$\vdots$	$\vdots$
$\pi_j =$	$a_{j+R} \pi_0 + a_{j+R-1} \pi_1 + \cdots + a_1 \pi_{j+R-1} + a_0 \pi_{j+R}$
$\vdots$	$\vdots$
$\pi_{J-1} =$	$a_{J+R-1} \pi_0 + a_{J+R-2} \pi_1 + \cdots + a_R \pi_{J-1} + a_{R-1} \pi_J$
$\pi_J =$	$a_{(J+R)^+} \pi_0 + a_{(J+R-1)^+} \pi_1 + \cdots + a_{(R+1)^+} \pi_{J-1} + a_{R^+} \pi_J$

completion of the data phase (and before the new packets are generated at the beginning of the next frame). Let X_t denote the number of packet (copies) in the queue in a given frame t. Let $\pi_j[t] = P[X_t = j],\ 0 \le j \le J$. (We impose a maximum queue occupancy J for calculation convenience and set it so large that boundary effects are negligible. Our analysis accommodates small J to study the effect of finite scheduling window and buffers in a straightforward manner.) The time-dependent probabilities $\pi_j[t]$ converge due to the ergodicity of the corresponding aperiodic Markov chain to the steady state probabilities, so we drop the time index t. The stationary probability π_0 for the state 0 is given by the first equation in Table I. To understand this equation, note that we return to $X_{t+1} = 0$ at the end of frame $t+1$ if we have no backlog at the end of the preceding frame ($X_t = 0$) and R or less packets arrive in frame $t+1$. If we have a backlog of one packet at the end of the preceding frame ($X_t = 1$), we return to $X_{t+1} = 0$ if $R-1$ or less packets arrive in frame $t+1$, etc. The set of linear equations for the other stationary probabilities π_j are obtained as given in Table I, where $a_{(c)^+} = P[A \ge c]$. Due to normalization

$$\sum_{j=0}^{J} \pi_j = 1. \tag{9}$$

We have thus a system of $J+2$ linear equations and $J+1$ unknowns. We apply Gaussian elimination with pivoting to numerically solve for the vector $[\pi_0, \pi_1, \ldots, \pi_J]$. The expected queue length $E[X]$ is given by

$$E[X] = \sum_{j=1}^{J} j \cdot \pi_j. \tag{10}$$

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

$$W_{TR} = \frac{E[X]}{S \cdot \sigma_q}. \tag{11}$$

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, see [22] for details. In brief, 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.

TABLE II

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

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, i.e., 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 $\kappa \cdot D$, 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

$$P(X_{[\delta]} = j) = \delta \cdot \left[\sum_{l=1}^{j} P(X = l) \right]^{\delta - 1} \cdot P(X = j). \tag{12}$$

Hence, approximately

$$\begin{aligned} E[\hat{X}] \quad = \quad & \sum_{\delta=1}^{\kappa \cdot D} \left(\sum_{j=1}^{J} j \cdot P(X_{[\delta]} = j) \right) \cdot P(\Delta = \delta) \\ & + E[X] \cdot \sum_{\delta = \kappa \cdot D + 1}^{D} P(\Delta = \delta), \end{aligned} \tag{13}$$

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

$$W_M = \frac{E[\hat{X}]}{S \cdot \sigma_q}. \tag{14}$$

V. PERFORMANCE RESULTS

In this section we numerically study 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. Initially, we fix the number of network nodes at $N = 200$ and the number of used wavelengths (transceivers at each node) at $\Lambda = 8$. The network parameters are summarized in Table II. 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. 8 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 $\Lambda = 8$

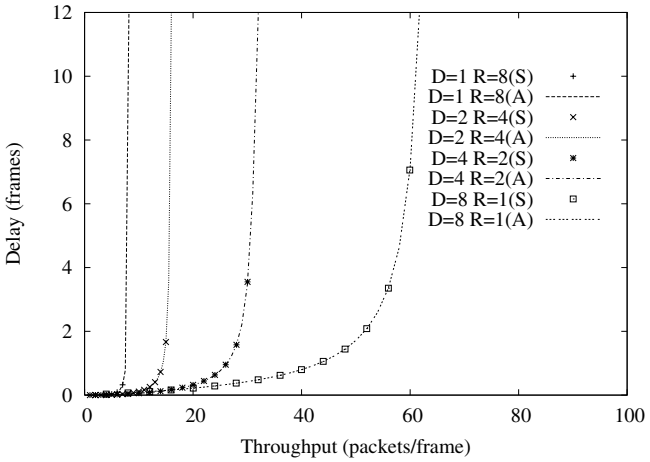

Fig. 8. Delay W_M as a function of throughput Z_M for unicast traffic $(u = 1)$.

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. Due to spatial wavelength reuse the total number of channels for the $(D = 8, R = 1)$ network is $D \cdot \Lambda = 64$. The maximum throughputs for the other three network configurations $(D = 4, R = 2)$, $(D = 2, R = 4)$, and $(D = 1, R = 8)$ are 32, 16, and 8 packets per frame, respectively.

B. Multicast Traffic

In Figures 9 and 10 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 (i) a maximum multicast throughput of 8 packets/frame, and (ii) the maximum receiver throughput of 800 packets/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, i.e., the number of completed multicasts per frame, is 8 packets/frame and the receiver throughput is 1600 packets/frame. Note that in this broadcast scenario the transmitter throughput is 8 packets/frame in the $(D = 1, R = 8)$ network and 64 packets/frame in the $(D = 8, R = 1)$ network.

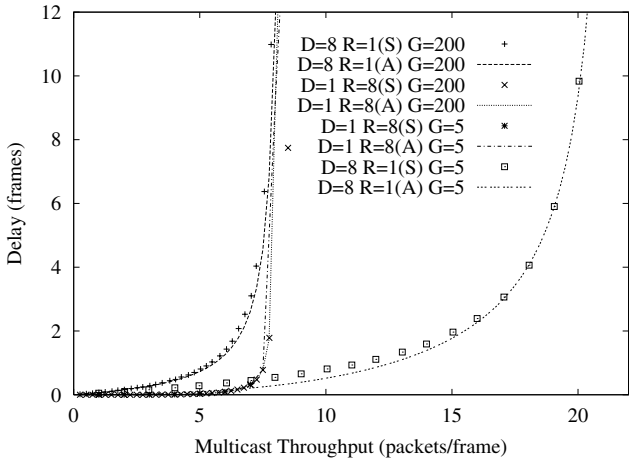

Fig. 9. Delay W_M as a function of multicast throughput Z_M for multicast traffic $(1 - u = 1)$ with $\Gamma = 5$ and $\Gamma = 200$.

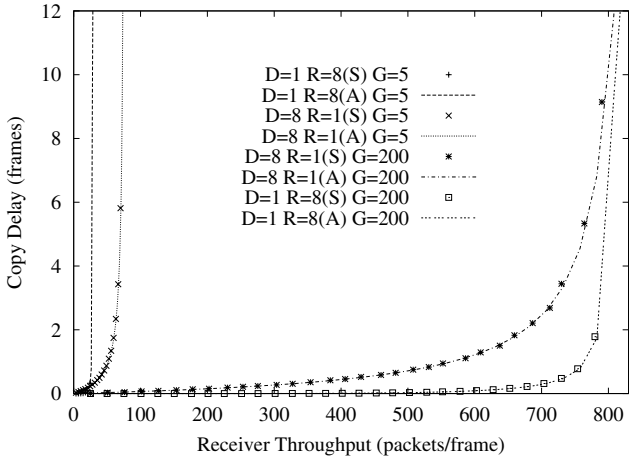

Fig. 10. Copy delay W_{TR} as a function of receiver throughput Z_R for multicast traffic $(u = 0)$ with $\Gamma = 5$ and $\Gamma = 200$.

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 $P(\Delta = D|\gamma = 100) = 0.98$. Thus it is very likely that D copies of the multicast packet need to be transmitted. In general, when multicasting over the $FT^{\Lambda} - FR^{\Lambda}$ 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) increasing 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, i.e., 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. 10, for a maximum multicast group size of $\Gamma = 5$ and a copy delay of 2 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 [17], [24], 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 [17], [24].

Returning to multicasting with transceivers arrays, which overcome the receiver bottleneck, we observe from Figures 9 and 10 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 summary, we find that the $FT^{\Lambda}-FR^{\Lambda}$ 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. 11 we plot the throughput–delay performance of the $FT^{\Lambda} - FR^{\Lambda}$ AWG network for 80% unicast traffic and 20% multicast traffic for different network configurations. We observe that with increasing AWG degree D the network achieves significantly larger multicast and receiver throughputs while the delay is increased only very slightly (at lower throughput levels). The throughput levels of the $(D = 8, R = 1)$ configuration are more than 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

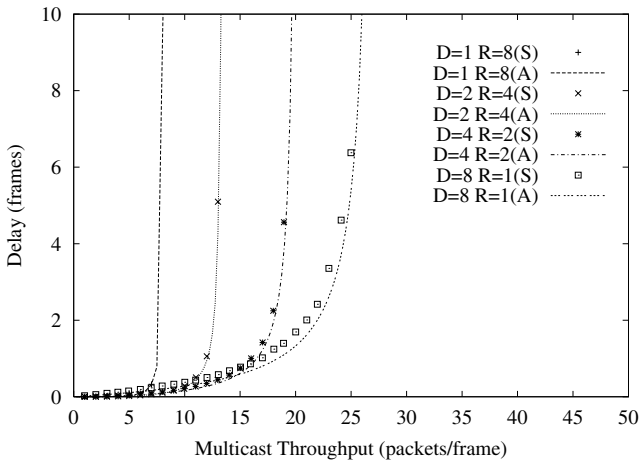

Fig. 11. Delay W_M as a function of multicast throughput Z_M for mix of 80% unicast $(u = 0.8)$ and 20% multicast traffic with $\Gamma = 200$.

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, i.e., 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.

In Figures 12 and 13, we plot the receiver throughput–delay performance for 60% and 90% unicast traffic. We observe

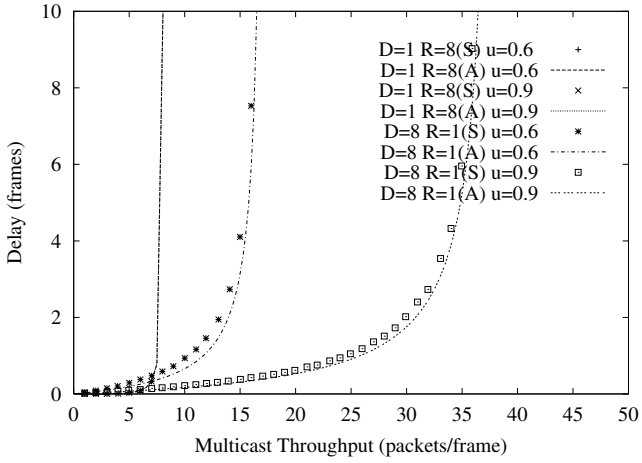

Fig. 12. Delay W_M as a function of multicast throughput Z_M for mixed traffic $u = 0.9$ and $u = 0.6$.

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 three times the throughput of the $(D = 1, R = 8)$ network; although the receiver throughput level is overall reduced for the larger portion of unicast traffic. Again we observe that the increase in throughput comes at the expense of only a minor increase in delay (nicely visible for the $u = 0.6$ scenario in

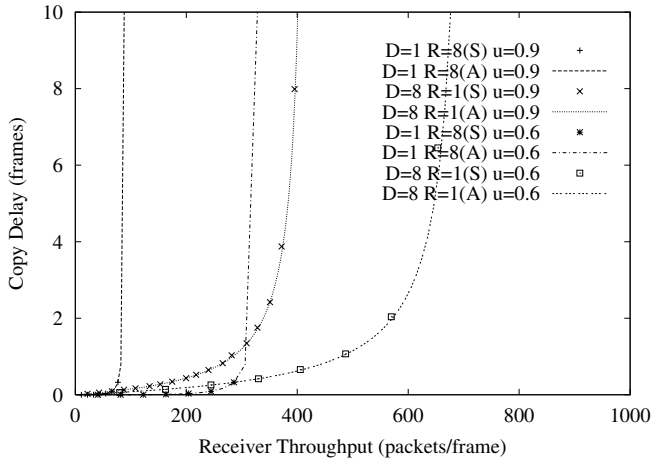

Fig. 13. Copy delay W_{TR} as a function of receiver throughput Z_R for mixed traffic $u = 0.9$ and $u = 0.6$.

TABLE III

THROUGHPUT DELAY FOR $(D = 8, R = 1)$ NETWORK FOR MULTICAST TRAFFIC $(u = 0.8)$ WITH $\Gamma = 200$

σ	Z_M	Z_T	Z_R	W_M	W_{TR}
0.01	2.0	4.6	38.6	0.06	0.04
0.02	4.0	9.4	79.7	0.12	0.08
0.04	8.0	18.8	160.4	0.28	0.21
0.08	15.9	37.4	319.8	0.86	0.67
0.125	25.0	58.7	501.9	6.38	5.28
0.135	27.0	63.9	549.0	150.5	127.3

the throughput range from $100 - 280$ packets/frame).

We observe that the accuracy of our probabilistic analysis is overall quite good. The discrepancies between the analytical and simulation results for the delay W_M for larger D are primarily due to the heuristic approximation (13) of the occupancy distribution of the longest queue, for which we set $\kappa = 0.75$ throughout this paper.

Tables III and IV show the detailed throughput–delay performance metrics obtained from simulation for the scenario with 80% unicast and 20% multicast traffic for the $(D = 1, R = 8)$ and $(D = 8, R = 1)$ network configurations. The stability limit (capacity) for the $(D = 8, R = 1)$ network is $D \cdot \Lambda = 64$ packets per frame. We observe that for a packet generation probability σ of 0.08 and less, corresponding to a transmitter throughput Z_T of 37.4 or less, or equivalently less than 58% of the capacity, the delays are very small. As the load increases to 90% of the capacity and higher, the delays become quite large. We also observe from the tables that for the $(D = 8, R = 1)$ network the average copy delay W_{TR} is for lower loads typically 75% or less of the corresponding delay

TABLE IV

THROUGHPUT DELAY FOR $(D = 1, R = 8)$ NETWORK FOR MULTICAST TRAFFIC $(u = 0.8)$ WITH $\Gamma = 200$

σ	Z_M	Z_T	Z_R	W_M	W_{TR}
0.01	2.0	2.0	41.8.0	0.0	0.0
0.02	4.0	4.0	83.7	0.01	0.01
0.035	7.0	7.0	146.5	0.33	0.33
0.040	0.79	8.0	167.5	45.4	45.4

TABLE V

MULTICAST THROUGHPUT Z_M FOR DELAY W_M OF 4 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 VI

RECEIVER THROUGHPUT Z_R FOR COPY DELAY W_{TR} OF 4 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

W_M for completing the transmission of all packet copies.

In Tables V and VI 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 V we fix the delay at 4 frames and record the maximum multicast throughput. In Table VI we fix the copy delay at 4 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. In general, 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 with a large average number of destination nodes. However, for mixed unicast and multicast traffic, both the multicast throughput and the receiver throughput improves significantly as D increases. Both the multicast throughput and the receiver throughput for the $(D = 8, R = 1)$ configuration are over 3 times that of the $(D = 1, R = 8)$ PSC network.

D. Impact of Number of Transceivers

In this section we study the throughput–delay performance of the $FT^{\Lambda} - FR^{\Lambda}$ AWG network for different numbers of transceivers Λ in each node. Throughout this section we fix the number of network nodes at $N = 200$ and the number of used FSRs at $R = 1$, hence $D = \Lambda$. Recall from Section II that the length of the control phase is N/Λ slots, each carrying one control packet. For our numerical evaluations in this and the following sections we consider a control packet length of 2 bytes and a data packet length of 1500 bytes. Thus the length of the control phase varies between 200 slots (for the degenerate case of) $\Lambda = 1$ and 25 slots for $\Lambda = 8$. The corresponding frame length varies between 950 slots and 775 slots. In Fig. 14 we plot the throughput–delay performance for the different Λ $(= D)$. The delay is given in slots and the throughput is given in steady state, i.e., normalized by the ratio of data phase to total frame length. We observe that the throughput for a fixed tolerable delay approximately triples as the number of nodal transceivers Λ doubles. There

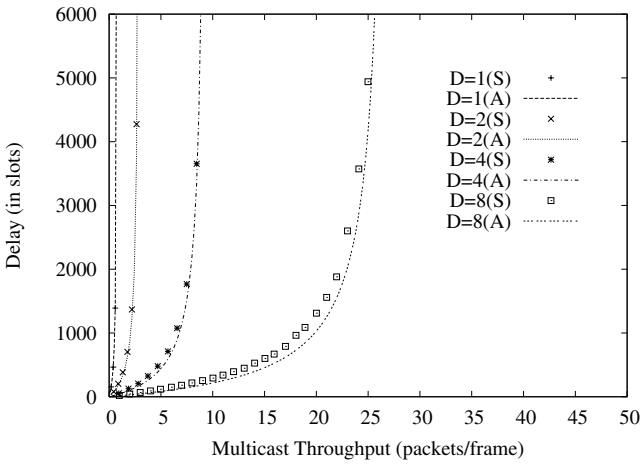

Fig. 14. Delay W_M as a function of multicast throughput Z_M for mix of 80% unicast ($u = 0.8$) and 20% multicast traffic with $\Gamma = 200$ for single FSR ($R = 1$) networks.

are two main effects at work here. On the one hand, the doubled number Λ of transceivers and the doubled wavelength reuse (governed by $D = \Lambda$) together quadruple the network capacity $D \cdot \Lambda$ (maximum number of data packet transmissions per frame). On the other hand, the increased number of required packet copy transmissions (for the larger D) results in increased delay. Overall, we observe that large throughputs are achieved for small numbers of transceivers Λ due to the extensive wavelength reuse on the AWG.

E. Control Packet Transmission: TDMA vs. contention

In this section we examine the impact of the TDMA and contention based control packet transmission strategies outlined in Sections III-A and III-B. We consider the $FT^{\Lambda} - FR^{\Lambda}$ AWG network with $D = 4$ and $R = 2$ for a mix of 80% unicast ($u = 0.8$) and 20% multicast traffic with $\Gamma = 200$. The length of the data phase is fixed at 1500 bytes or equivalently 750 slots throughout. In Fig. 15 we compare the throughput–delay obtained from simulation for (i) TDMA control packet transmission with a control phase with N/Λ slots, and (ii) control packet transmission with contention with a control phase with M = 5 and 10 slots. We observe from Fig. 15a) that for N = 200 nodes, TDMA control packet transmission gives better throughput delay performance than control packet contention. This is because the effect of the slightly shorter control phase with contention is outweighed by the delay introduced due to control packet collisions and subsequent retransmissions. Note that each retransmission introduces an additional delay of one frame, whereby in the considered scenario the frame is significantly longer than the control phase.

Control packet transmission with contention is advantageous when the length N/Λ of the TDMA control phase makes up a significant portion of the frame length, i.e., when either the number of nodes N is large or the data packets are short. We illustrate this effect by scaling up the number of nodes to N = 2000 in Fig. 15b). We observe that in this scenario, control packet contention with a data phase consisting of M = 10 slots gives consistently better throughput-delay performance than TDMA control packet transmission. This is because in this scenario, the effect of the significantly shorter control phase with contention outweighs the effect of occasional control packet collisions and retransmissions. If the number of control slots is too small, then the control packet contention becomes increasingly a bottleneck as the traffic load increases, as illustrated in Fig. 15b) for M = 5.

F. 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 (i) a TT–TR AWG network where the control packets are transmitted with an LED (as in [2]) over the AWG, and (ii) 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.

In Figure 16, we plot the throughput–delay performances of

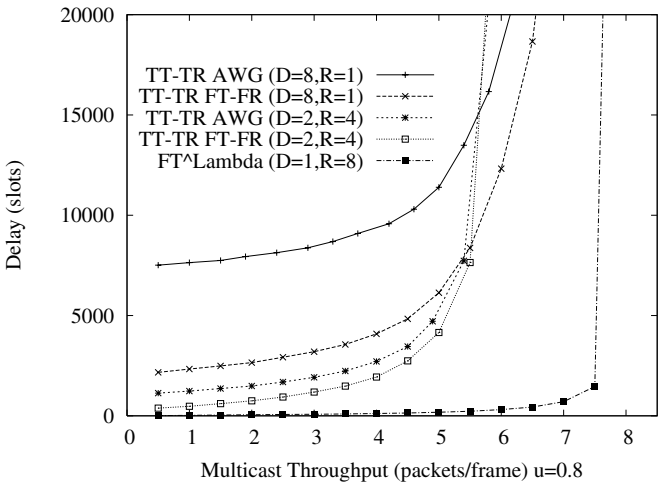

Fig. 16. Delay W_M as a function of multicast throughput Z_M for TT–TR AWG, TT–TR–FT–FR AWG, and $FT^{\Lambda} - FR^{\Lambda}$ AWG networks.

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. 11. We observe that 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, has significantly lower performance than the worst performing $FT^{\Lambda} - FR^{\Lambda}$ AWG network configuration.

VI. Conclusion

In this paper we have developed and evaluated the $FT^{\Lambda} - FR^{\Lambda}$ AWG network, an AWG based single–hop metro WDM

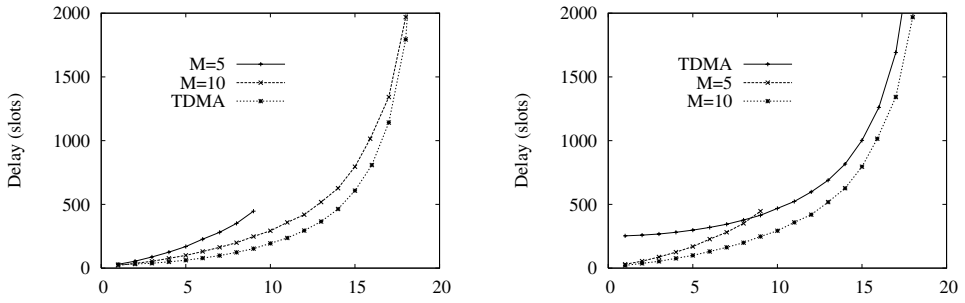

Fig. 15. Delay W_M in slots as a function of multicast throughput Z_M for control packet transmission with TDMA and contention.

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. Our analytical and simulation results 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. In our ongoing work we are studying efficient protection strategies for the network.

ACKNOWLEDGMENT

This work was supported in part by National Science Foundation through grant Career ANI–0133252 and the DFG research center Mathematics for Key Technologies (FZT86) in Berlin. We are grateful to Martin Maier for insightful discussions during the early stages of this research.

REFERENCES

[1] B. Mukherjee, "WDM optical communication networks: Progress and challenges," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 1810–1824, Oct. 2000.

[2] M. Scheutzow, M. Maier, M. Reisslein, and A. Wolisz, "Wavelength reuse for efficient packet-switched transport in an AWG-based metro WDM network," *IEEE/OSA Journal of Lightwave Technology*, vol. 21, no. 6, pp. 1435–1455, June 2003.

[3] 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. of OFC 2001*, vol. 4, Anaheim, CA, Mar. 2001, pp. ThG–1–ThG–3.

[4] 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 Photonics Technology Letters*, vol. 12, pp. 1549–1551, November 2000.

[5] 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 Photonics Technology Letters*, vol. 11, pp. 1114–1116, September 1999.

[6] M. Maier, M. Reisslein, and A. Wolisz, "Towards efficient packet switching metro WDM networks," *Optical Networks Magazine*, vol. 3, no. 6, pp. 44–62, November/December 2002.

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

[8] A. Hamad and A. Kamal, "A survey of multicasting protocols for broadcast–and-select single–hop networks," *IEEE Network*, vol. 16, pp. 36–48, July/August 2002.

[9] J. He, S.-H. G. Chan, and D. H. K. Tsang, "Multicasting in WDM networks," *IEEE Communications Surveys and Tutorials*, Dec. 2002.

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

[11] L. Wang and K. Lee, "A WDM based virtual bus for universal commuincation and computing systems," in *Proc. of ICC '92*, June 1992, pp. 888–894.

[12] K. Kato, A. Okada, Y. Sakai, and K. N. *et al.*, "10–Tbps full–mesh WDM network based on a cyclic–frequency arrayed–waveguide grating router," in *Proc. of ECOC '00*, vol. 1, Munich, Germany, Sept. 2000, pp. 105–107.

[13] A. Okada, T. Sakamoto, Y. Sakai, and K. N. *et al.*, "All–optical packet routing by an out–of–band optical label and wavelength conversion in a full–mesh network based on a cyclic–frequency AWG," in *Proc. of OFC 2001 Technical Digest, paper ThG5*, Anaheim, CA, Mar. 2001.

[14] D. Banerjee, J. Frank, and B. Mukherjee, "Passive optical network architecture based on waveguide grating routers," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 7, pp. 1040–1050, Sept. 1998.

[15] A. Bianco, E. Leonardi, M. Mellia, and F. Neri, "Network controller design for SONATA — a large–scale all–optical passive network," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 2017–2028, Oct. 2000.

[16] N. P. Caponio, A. M. Hill, F. Neri, and R. Sabella, "Single–layer optical platform based on WDM/TDM multiple access for large–scale 'switchless' networks," *European Transactions on Telecommunications*, vol. 11, no. 1, pp. 73–82, Jan./Feb. 2000.

[17] M. Maier, M. Scheutzow, and M. Reisslein, "The arrayed–waveguide grating based single–hop WDM network: an architecture for efficient multicasting," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 9, pp. 1414–1432, Nov. 2002.

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

[19] 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 Journal on Selected Areas in Communications*, vol. 16, no. 7, pp. 1134–1145, Sept. 1998.

[20] Y. Sakai, K.Noguchi, R. Yoshimura, and T. S. *et al.*, "Management system for full-mesh WDM AWG-STAR network," in *Proc. of ECOC '01.*, vol. 3, Amsterdam, Netherlands, Sept. 2001, pp. 264 –265.

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

[22] 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," Arizona State University, Tech. Rep., Dec. 2003, available at `http://www.fulton.asu.edu/~mre`.

[23] V. Kolchin, B. Sevastyanov, and V. Chistyakov, *Random Allocations*. Wiley, 1978.

[24] H. C. Lin and C. H. Wan, "A hybrid multicast scheduling algorithm for single–hop WDM networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 19, no. 11, pp. 1654–1664, Nov. 2001.