Performance Analysis of Optical Flow Switching

Guy Weichenberg, Student Member, IEEE, Vincent W. S. Chan, Fellow, IEEE, Muriel Médard, Fellow, IEEE
Research Laboratory of Electronics, Massachusetts Institute of Technology

Abstract—In our previous work [1], [2], we presented Optical Flow Switching (OFS) as a key enabler of scalable future optical networks. In the present work, we propose a practical scheduling algorithm and conduct an approximate throughput-delay analysis for OFS networks.

I. INTRODUCTION

In the four decades since optical fiber was introduced as a communications medium, optical networking has revolutionized the telecommunications landscape. Sustained exponential growth in communications bandwidth demand, however, is requiring that the nexus of innovation in optical networking continue, in order to ensure cost-effective communications in the future. In our previous work [1], [2], we presented OFS as a key enabler of scalable future optical networks. In the present work, we propose a scheduling algorithm to arbitrate access to resources in OFS networks in a manner that only loosely couples pairs of communicating metropolitan-area networks (MANs). We then conduct an approximate throughput-delay analysis of OFS networks. In the next section, we outline the physical layer and traffic assumptions employed throughout this work. In section IV, we discuss the properties of a sensible scheduling algorithm for inter-MAN OFS communication, and then proceed to propose our own practical algorithm which meets these criteria. In section V, we derive an analytic approximation of the throughput-delay tradeoff for OFS under our scheduling algorithm. In section VI, we carry out a numerical study of our results in section V to further our understanding of the tradeoffs in the OFS architecture design space. We conclude this work in section VII.

II. OFS OVERVIEW

Owing to space constraints, we present only a high-level overview of OF and refer the interested reader to [3] for further details. In OF, users request end-to-end lightpaths for long duration (i.e., greater than 100 ms) transactions. In order to schedule data transmission across the wide-area network (WAN), users communicate via an electronic control plane with the scheduling processors assigned to their respective MANs. These scheduling processors, in turn, coordinate transmission of data across the WAN in an electronic control plane. This is in contrast to Optical Burst Switching (OBS), where access to network resources occurs in a random-access fashion. In OFS, it is assumed that the smallest granularity of bandwidth that can be reserved across the core is a wavelength. In the event that several single users have transactions which are not sufficiently large to warrant their own wavelength channels, they may multiplex their data for transmission across the WAN via dynamic broadcast group formation.

Motivated by the minimization of network management and switch complexity in the network core, transactions are serviced as indivisible entities. That is, data cells comprising a flow traverse the network contiguously in time, along the same wavelength channel (assuming no wavelength conversion), and along the same spatial network path. This is in contrast to packet switched networks, where transactions are broken up into constituent cells, and these cells are switched and routed through the network independently. Note that in OFS networks, unlike packet switched networks, all queuing of data occurs at the end users, thereby obviating the need for buffering in the network core. A core node is thus equipped with a bufferless optical cross-connect (OXC). OFS is a centralized transport architecture in that coordination is required for logical topology reconfiguration. However, OFS traffic in the core will likely be sufficiently aggregated and intense to warrant a quasi-static logical topology that changes on coarse time scales.

III. MODELING ASSUMPTIONS

A. Network topology and other physical layer issues

In our network model, a single WAN connects \( n_w \) MANs, all of which employ OFS. A MAN is connected to the WAN via a single MAN node residing at the MAN-WAN interface. The wavelength channels provisioned for inter-MAN OFS communication reside within \( f \) fibers in each direction connecting this node to the rest of the WAN.

We assume MAN physical topologies to be based upon arbitrary mesh designs. An OFS MAN node comprises an OXC with direct connections to adjacent MAN nodes as well as one or more access networks based upon the all-optical distribution network (DN) architectures. We denote the total number of such DNs per MAN by \( n_d \). Under normal operating conditions, inter-MAN traffic is assumed to be carried along an embedded tree of the MAN physical topology, as drawn in Figure 1; whereas the portion of the topology outside the embedded tree (not drawn in the figure) carries only intra-MAN traffic. Since intra- and inter-MAN OFS traffic could coexist on the same fiber in the embedded tree, we will assume that \( w_1 \) channels are (quasi-statically) allocated for inter-MAN OFS communication. This separation between wavelength channels for inter- and intra-MAN communication is made for analytical tractability, but may also prove to be sensible in implementing real networks since it enables simpler resource scheduling decisions, albeit with some performance penalty.

For reasons of scheduling algorithm simplicity, we require that, for each WAN wavelength channel provisioned for inter-MAN OFS communication, there exists a dedicated wavelength channel in each link of the embedded tree in both source and destination MANs. In the absence of wavelength conversion capability at the MAN-WAN interface, wavelength continuity between WAN and MAN wavelength channels must be respected.

978-1-4244-3435-0/09/$25.00 ©2009 IEEE
The exact number of required fibers per link to satisfy this is situation dependent; so we shall assume for the remainder of this work that the links in the embedded tree topology include $f$ fibers in each direction to ensure the one-to-one correspondence to the fibers connecting the MAN to the WAN.

With respect to the number of fibers connecting each DN to its parent MAN node, we consider two extreme cases: i) two fibers, one in each direction, and ii) $2f$ fibers, $f$ in each direction such that there is one-to-one correspondence to the $2f$ fibers connecting the MAN to the WAN. Clearly, the performance corresponding to the second case will be better; but, as we shall see, the performance margin is not great under expected future network dimensions. Moreover, the case of two fibers per DN is attractive in that it is a simpler design: less hardware at end-users is required, and no modifications to this hardware are required as the number of fibers $f$ in the MAN increases. We also consider the role of wavelength conversion between each DN and its parent MAN node.

### B. Traffic

We will confine our attention in this work to serving the inter-MAN traffic demand. We will assume uniformity in inter-MAN traffic demands, as well as uniformity in DN traffic demands. The former assumption implies that each MAN communicates with every other MAN on at most $w_m = \frac{\lambda_m}{n_p-1}$ wavelength channels in each direction.

OFS traffic is assumed to be generated at end-users such that the aggregate traffic generated for a particular destination MAN arrives according to a Poisson process with rate $\lambda_{in}$. This Poisson assumption is reasonable for commercial networks, since the superposition of many stationary, identical, and independent point processes—which are reasonable models themselves for individual flow sources—is well-known to converge to a Poisson point process. In the event that there do not exist sufficiently many flow sources to multiplex, we argue that OFS may not be an appropriate architecture.

The duration of flows are modeled as identical and independently distributed random variables with probability density function $p_L(\cdot)$ and $k$th moment $\overline{T_F}$. Lastly, we consider only unicast transactions—in part, because of the absence of commercially viable OXCs capable of multicast—although we recognize the increasing importance of multicast transactions, particularly for video content distribution.

### IV. SCHEDULING FOR INTER-MAN COMMUNICATION

Before outlining the scheduling algorithm used to serve inter-MAN traffic, we review features of OFS networks that provide guidelines for our scheduling algorithm design.

As discussed in the previous section, we focus on networks for which there exists significant multiplexing of flows in each MAN. Following the work of Cao et al. in [4], appreciable statistical smoothing arises from the multiplexing of many flows on wavelength channels. Such smoothing of aggregated traffic renders a quasi-static WAN logical topology sensible for serving this traffic. That is, changes to the WAN logical topology will be required on time-scales that are on the order of many flows. Consequently, in the scheduling algorithm design for the scheduling of individual flows it may be assumed that the wavelength channels provisioned for inter-MAN communication are static. This, along with our method of provisioning wavelength channels in the MAN, allows us to significantly decouple the scheduling of resources among pairs of communicating MANS.

Owing to the high cost of provisioning wavelength channels in the wide-area, a scheduled approach should be employed which ensures high utilization of these resources. Ideally, this scheduled approach would be enable reservation of end-to-end optical paths without decoupling of the three geographic network tiers. However, such approaches are implementationally infeasible owing to the computational effort given the scale of the problem. Even polynomial-time heuristics, such as linear-programming relaxations, require prohibitive computational effort given the size of an end-to-end OFS scheduling problem. Indeed, in OFS, the number of resources to be scheduled is on the order of the product of the number of wavelength channels per fiber and the number of DNs sharing a WAN, which could easily be on the order of $10^5$. Running such algorithms to schedule individual flows would thus be infeasible.

Our approach in this section is to exploit the unique characteristics of the different geographic network tiers in OFS to devise an algorithm which, in addition to having constant complexity, lends to an analytic delay study. Our simple scheduling algorithm, in particular, reserves end-to-end optical paths for flow transmission via a sequential reservation process in which wavelength channels in the MAN and WAN are reserved first, followed by simultaneous but separate reservations of the source and destination DN wavelength channels. Owing to the fact that wavelength channels in DNs are not heavily loaded, the latter reservation step should entail very little contention, thereby ensuring efficient end-to-end scheduling.

#### A. Scheduling algorithm description

In the following description of the scheduling algorithm, we assume that each DN is connected to its MAN by two fibers. The case of $2f$ fibers is simpler and will be addressed at the end. In addition, we neglect the possibility of transmitter and receiver collisions which arise when two or more flows simultaneously require an end-user’s transmitter or receiver, respectively. This is a reasonable assumption when each end-user transmits flows only occasionally, which we assume to be the case. We now
illustrate the scheduling algorithm by stepping through the process by which an end-to-end all-optical path is established for the transmission of a particular flow.

Consider a flow generated at an end-user residing in DN $D_s$ within MAN $M_s$ and that is destined for an end-user residing in DN $D_d$ within MAN $M_d$. As soon as this flow is ready for transmission, the source end-user sends “primary request” $r_w$ to the scheduling node associated with $M_s$ requesting an end-to-end all-optical path for its flow transmission.

At a MAN’s scheduling node, there exist $n_w - 1$ first-in-first-out (FIFO) queues, one queue corresponding to every possible MAN destination. Each queue can be thought as the queue for an $M/G/w_m$ queueing system, in that the $w_m$ wavelength channels dedicated to transmission from $M_s$ to $M_d$ eventually serve the primary requests waiting in it. After it arrives at $M_s$’s scheduling node, $r_w$ is placed at the head of the queue associated with $M_d$ which we denote $Q_{M_d}^{M_s}$. Once $r_w$ reaches the head of $Q_{M_d}^{M_s}$, $r_w$’s flow is assigned to wavelength channel $\omega$, the first of the $w_m$ wavelength channels dedicated to transmission from $M_s$ to $M_d$ to have the primary request it is serving depart. After this wavelength channel assignment is made, an all-optical path is established on wavelength channel $\omega$ from the edge of $D_s$ to the edge of $D_d$, passing through $M_s$, the WAN, and $M_d$. (Such a path is guaranteed to exist since there are 2 $f$ fibers within each link on this path, one of which with a dedicated $\omega$ wavelength channel for communication from $M_s$ to $M_d$.) Now, in order to reserve the single outgoing $\omega$ wavelength channel in $D_s$ and the single incoming $\omega$ wavelength channel in $D_d$, two additional secondary requests, $r_s$ and $r_d$, respectively, are sent. During this process, $r_w$ remains at the head of $Q_{M_d}^{M_s}$.

Secondary request $r_s$ joins the end of the “source” secondary queue associated with $D_s$’s $\omega$ wavelength channel, denoted by $Q_{M_s}^{D_s}(\omega)$, which is physically located in $M_s$’s scheduling node; and secondary request $r_d$ joins the end of the “destination” secondary queue associated with $D_d$’s $\omega$ wavelength channel, denoted by $Q_{M_d}^{D_d}(\omega)$, that is physically located in $M_d$’s scheduling node. These queues contain secondary requests to use the $\omega$ wavelength channel on $D_s$’s outgoing fiber and $D_d$’s incoming fiber, respectively. Note that at any instant in time there can be at most one secondary request associated with flows destined for $M_d$. When $r_s$ and $r_d$ each reach the heads of their respective secondary queues (which we assume to be FIFO), they each notify both $M_s$’s and $M_d$’s’ scheduling nodes. As soon as $M_s$’s and $M_d$’s’ scheduling nodes have received both notifications, they instruct the source and destination end-users, respectively, to begin flow transmission immediately. After the flow transmission is complete, $r_w$, $r_s$, and $r_d$ depart their queues.

The scheduling algorithm for the case in which a DN is connected to its MAN with 2 $f$ fibers is simpler because the source and destination DN queues $Q_{M_s}^{D_s}(\omega)$ and $Q_{M_d}^{D_d}(\omega)$ are no longer necessary. This is because the correspondence between each of these fibers and the 2 $f$ fibers in the MAN and WAN preclude contention for DN wavelength channel resources among flows with different destination MANs. Thus, for the case of 2 $f$ fibers, once primary request $r_w$ reaches the head of $Q_{M_d}^{M_s}$, instruction is sent to the scheduling node associated with $M_d$ to set up an all-optical path in $M_d$ from the WAN to $D_d$ on channel $\omega$. The two scheduling nodes then instruct the end-users to begin flow transmission immediately.

V. APPROXIMATE PERFORMANCE ANALYSIS

In the following performance analysis, we neglect propagation delay of transactions (and requests). This is a reasonable assumption since the OFS architecture has been shown to be most appropriate for serving large transactions [1].

In our description of the scheduling algorithm for inter-MAN communication in section IV-A, we mentioned that primary requests for a source-destination MAN pair may be modeled as customers of an $M/G/w_m$ queueing system with arrival rate $\lambda_m$. The service time in this model is the time spent by a primary request $r_w$ at the head of its primary queue, which comprises flow transmission time in addition to the time spent reserving wavelength channels in the source and destination DN (in the case of two fibers per DN). Unfortunately, there is no exact solution for the $M/G/w_m$, except in the special cases of exponential service times, $w_m = 1$, or $w_m = \infty$. We thus resort to an approximation of the performance of the $M/G/w_m$ queueing system based upon the single server queueing system model of our OFS network [5].

To this end, we randomly split the Poisson process of intensity $\lambda_m$ representing the flow arrivals of each source-destination MAN pair into $w_m$ Poisson processes of intensity $\lambda_c = \lambda_m / w_m$, one baby Poisson arrival process for each wavelength channel dedicated to the source-destination MAN pair. We also replace queue $Q_{M_d}^{M_s}$ in MAN $M_s$ by $w_m$ independent, parallel queues $Q_{M_d}^{M_s}(\omega_1), Q_{M_d}^{M_s}(\omega_2), \ldots, Q_{M_d}^{M_s}(\omega_{w_m})$, each corresponding to a wavelength channel dedicated to the source-destination MAN pair and accepting primary requests from the corresponding baby Poisson process.

To compute the queuing delay of this system we will eventually apply the Pollaczek-Khinchin (P-K) formula (e.g., see [6]) to the single server queue $Q_{M_d}^{M_s}(\omega)$ accepting primary requests from flows generated in $M_s$, destined for $M_d$, and employing wavelength channel $\omega$. The arrival rate to this queue is $\lambda_c$, and we shall denote the service time of a primary request in this queue by $X$ with $k$th moment $E[X^k]$. In the case of 2 $f$ fibers per DN, we simply have $X = L + \tau$, where $\tau$ is the hardware reconfiguration time, since no additional time is required to reserve resources at the source and destination DNs. However, in the case of two fibers per DN, $X$ comprises not only the sum of the flow transmission time $L$ and hardware configuration time $\tau$, but also any time spent at the head of queue $Q_{M_d}^{M_s}(\omega)$ reserving source and destination DN resources. We therefore generally express the average service time $\overline{X}$ as:

$$\overline{X} = \overline{L} + \overline{\tau} \approx \overline{L} + \overline{\tau}$$

where $\overline{\tau}$ is the average queuing time spent at the head of the primary queue and is equal to zero in the case of 2 $f$ fibers per DN. The approximation for the average service time neglects
the hardware reconfiguration time, as \( L + \bar{Y} \) is, at the very least, on the order of hundreds of milliseconds, whereas \( \tau \) is likely to be on the order of a few tens of milliseconds.

Recall that, in the case of two fibers per DN, after a primary request reaches the head of \( Q_{M_s}^L(\omega) \), secondary requests are sent to each of \( \bar{Q}_{M_s}^{D_s}(\omega) \) and \( \bar{Q}_{M_s}^{D_d}(\omega) \), where \( D_s \) and \( D_d \) are the source and destination DNs of the flow, respectively. At each of these two secondary queues there could be up to \( f - 1 \) other secondary requests associated with other destination and source MANs, respectively. However, since flows are equally likely to be generated at, or destined for each of the DNs in a MAN, the probability that a primary request at \( Q_{M_s}^L(\omega) \) generates a secondary request to a particular \( \bar{Q}_{M_s}^{D_s}(\omega) \) or \( \bar{Q}_{M_s}^{D_d}(\omega) \) is \( 1/\bar{n}_a \). The arrival rate of secondary requests to a secondary queue contributed from each of the \( f \) contending primary queues is \( \lambda_c/\bar{n}_a \), for an aggregate arrival rate at each secondary queue of \( f\lambda_c/\bar{n}_a \ll 1 \). Now, for a fixed aggregate arrival rate of \( \lambda_c/\bar{n}_a \), as \( \bar{n}_a \) and \( \bar{n}_d \) get proportionately large, the arrival process of secondary requests to each secondary queue is known to converge to a Poisson process. We thus model the arrival process to each secondary queue as a Poisson process of rate \( f\lambda_c/\bar{n}_a \), with the approximation becoming increasingly accurate as \( \bar{n}_a \) and \( \bar{n}_d \) become large.

We now address the calculation of the moments of \( Y \). Recall that \( Y \) is the maximum of the two queueing delays experienced by the two peer secondary requests generated by a primary request once it reaches the head of its primary queue. Since the service time of each secondary request already enqueued at a secondary queue is itself coupled to the state of the queue in which its peer secondary resides, the characterization of \( Y \) is quite difficult. Thus, given the Poisson assumption of secondary request arrivals, we derive upper and lower bounds for \( \bar{Y} \) and \( \bar{Y}^2 \) which will be used in conjunction with the P-K formula to obtain optimistic and pessimistic approximations, respectively, for the queueing delay experienced by a flow. These approximations do not serve as strict bounds since the Poisson assumption is an approximation.

A. Optimistic approximation for primary request service time

We may compute simple lower bounds for \( \bar{Y} \) and \( \bar{Y}^2 \) by viewing the time spent to reserve the source and destination DNs as equal to the time spent by a single secondary request in a single queue with service time drawn from \( p_L(\cdot) \). This bound can be thought of as arising if only one of the two DNs needs to be reserved rather than both. In this case, the P-K formula yields the following bound for the first moment of \( Y \):

\[
\bar{Y} > \bar{Y}_L = \frac{f\lambda_c\bar{L}^2}{2(\bar{n}_a - f\lambda_c\bar{L})},
\]

and the following bound for the second moment of \( Y \):

\[
\bar{Y}^2 > \bar{Y}_L^2 = \frac{(f\lambda_c\bar{L}^2)^2}{2(\bar{n}_a - f\lambda_c\bar{L})^2} + \frac{f\lambda_c\bar{L}^3}{3(\bar{n}_a - f\lambda_c\bar{L})},
\]

after invoking the well-known Takács recurrence formula (e.g., see [6]) for the moments of the waiting time in an \( M/G/1 \) queueing system. These bounds can now be employed to yield the following approximations for the first two moments of the service time in the primary request queue:

\[
\bar{X} \approx \bar{L} + \bar{Y}_u = \bar{L} + \frac{f\lambda_c\bar{L}^2}{2(\bar{n}_a - f\lambda_c\bar{L})},
\]

\[
\bar{X}^2 \approx \bar{L}^2 + \bar{Y}_u^2 = \bar{L}^2 + \frac{(f\lambda_c\bar{L}^2)^2}{2(\bar{n}_a - f\lambda_c\bar{L})^2} + \frac{f\lambda_c\bar{L}^3}{3(\bar{n}_a - f\lambda_c\bar{L})},
\]

B. Pessimistic approximation for primary request service time

To compute upper bounds for \( \bar{Y} \) and \( \bar{Y}^2 \), we consider secondary requests, instead of being sent simultaneously after a primary request reaches the head of its queue, are sent sequentially. Specifically, we assume that the secondary request reserving the destination DN is sent only after the secondary request responsible for reserving the source DN reaches the head of its queue in effect reserving the source DN. Thus:

\[
\bar{Y} < \bar{Y}_u = \bar{Z}_s + \bar{Z}_d,
\]

where \( Z_s \) is the time spent by a secondary source request in its queue prior to reaching the head of the queue; and \( Z_d \) is the time spent by a secondary destination request in its queue prior to reaching the head of its queue.

By invoking the Poisson approximation for the arrival process to the secondary destination queue, we may thus treat the queue as an \( M/G/1 \) queueing system. The first three moments of \( Z_d \) may be obtained in a straightforward manner from the Takács recurrence formula. (We omit their explicit expressions for brevity.) To compute the first two moments of \( Z_s \), we invoke the Poisson approximation for the arrival process of secondary requests to the secondary source queue, and thus treat the queue as an \( M/G/1 \) queueing system. Note that in this queueing system, the service time of a customer is \( L + Z_d \). The first moment of \( Z_s \) is then:

\[
\bar{Z}_s = \frac{f\lambda_c(\bar{L}^2 + 2\bar{L} \cdot \bar{Z}_d + \bar{Z}_d^2)}{2(\bar{n}_a - f\lambda_c\bar{L} - f\lambda_c\bar{Z}_d)},
\]

and the following for the second moment of \( Z_s \):

\[
\bar{Z}_s^2 = \frac{1}{2} \left( \frac{f\lambda_c(\bar{L}^2 + 2\bar{L} \cdot \bar{Z}_d + \bar{Z}_d^2)}{\bar{n}_a - f\lambda_c\bar{L} - f\lambda_c\bar{Z}_d} \right)^2 \left( \frac{f\lambda_c(\bar{L}^2 + 2\bar{L} \cdot \bar{Z}_d + \bar{Z}_d^2)}{\bar{n}_a - f\lambda_c\bar{L} - f\lambda_c\bar{Z}_d} \right) + \frac{f\lambda_c(\bar{L}^2 + 3\bar{L}^2 \cdot \bar{Z}_d + 3\bar{L} \cdot \bar{Z}_d^2 + \bar{Z}_d^3)}{3(\bar{n}_a - f\lambda_c\bar{L} - f\lambda_c\bar{Z}_d)}.
\]

where the aforementioned expressions for the moments of \( Z_d \) should be substituted in.

We are now able to form the following pessimistic approximations for the first two moments of the service time in the primary request queue:

\[
\bar{X} \approx \bar{L} + \bar{Y}_u = \bar{L} + \bar{Z}_s + \bar{Z}_d
\]

\[
\bar{X}^2 \approx \bar{L}^2 + \bar{Y}_u^2 = \bar{L}^2 + \bar{Z}_s^2 + \bar{Z}_d^2 + 2\bar{L} \cdot \bar{Z}_s + 2\bar{L} \cdot \bar{Z}_d + 2\bar{Z}_s \cdot \bar{Z}_d.
\]
C. Wavelength conversion in DNs

In the presence of wavelength conversion capability in each DN, there is only one secondary source queue and one secondary destination queue associated with each DN. Each of these queues is served by \( w_i \) wavelengths channels akin to an \( M/G/1 \) queueing system. In this case, the maximum number of contending secondary requests at a queue would be \((n_w - 1)w_m\) rather than \( f \) as in the case of no wavelength conversion. However, since there are \( w_i \) candidate wavelength channels serving this queue instead of just one, the normalized arrival rate of secondary requests to a secondary queue is again \( f\lambda_c/n_a \approx 1 \). The convergence of this arrival process to Poisson is faster in this case than in the case of no wavelength conversion owing to the fact that \((n_w - 1)w_m > f \).

Furthermore, the delay experienced by secondary requests at these secondary queues will be less than in the case of no wavelength conversion owing to the statistical smoothing gain.

To quantitatively evaluate the approximate performance gain provided by wavelength conversion capability in DNs, we invoke a well-known, simple approximation of the delay experienced by a flow, including the time spent reserving DN wavelength channels; service time. Here, simply scale prior to reaching the head of the queue [5]. In particular, we invoke a well-known, simple approximation of the delay experienced by a primary request while at the head of its queue with an additional term reflecting the queueing delay of these queues is served by \( w \) wavelengths channels, respectively, as functions of offered load and service distribution.

D. Total queueing delay

Given the above optimistic and pessimistic approximations for the first two moments of \( X \), we now turn to computing the total queueing delay seen by a flow. To do this, we first invoke the P-K formula with respect to the primary request queue with an additional term reflecting the queueing delay experienced by a primary request while at the head of its queue. We then invoke the previous \( M/G/k \) approximation to obtain the following approximation of the total queueing delay experienced by a flow, including the time spent reserving DN wavelength channels just prior to flow transmission:

\[
W \approx \left[ Y + \frac{\lambda_w X^2}{2 \left(1 - \lambda_s X\right)} \right] \frac{W_{M,w,m}(\lambda_m X, p_X)}{W_{M,1}(\lambda_s X, p_X)}.
\]  

Equations (1) and (2), or equations (3) and (4), or their wavelength conversion counterparts may be substituted into equation (5) ultimately yielding optimistic and pessimistic approximations of \( W \), respectively.

VI. NUMERICAL RESULTS

In Figure 2, the approximations of the queueing delay derived from equation (5) are plotted versus WAN wavelength channel utilization (equation (6)) for three flow length distributions: constant, exponential, and truncated heavy-tailed, respectively. For each distribution, the approximations, which differ only in the way DN reservations are carried out, yield similar performances at low loads since there is little contention for DN resources. The performances diverge with increasing traffic load—especially between the case of no wavelength conversion (i.e., red curves) and the other cases—owing to the increasing role of DN reservation time. In comparing performance across distributions, we observe that constant length flows offer the best delay-throughput tradeoff, while truncated heavy-tailed length flows offer the worst. These results indicate very large flows impeding subsequent flows to the greatest extent in the truncated heavy-tailed distribution and to the least extent in the constant distribution.
In Figure 3, we illustrate the impact of the number of DNs per MAN on the delay-throughput tradeoff presented by equations (5) and (6). As expected, the performance of the two fibers per DN case converges to that of the 2f fibers per DN case as the number of DNs per MAN increases. This, of course, is because as ˜n increases, the amount of traffic per DN decreases, ultimately resulting in less contention for DN resources. We also observe the expected result that the gap between the optimistic and pessimistic approximations for a the same ˜n narrows as the number of DNs per MAN increases.

Since ˜n is a design parameter that is under the control of the network architect, it is interesting to investigate its relationship, under the assumption of two fibers per DN, with another notable design parameter: w, the number of wavelength channels provisioned for each MAN pair. Figure 4 depicts typical tradeoff curves for these two parameters juxtaposed with a simple inverse proportionality relationship for fixed first two moments of X. As illustrated, a scaling of ˜n by a factor of, say k, results in a scaling down of w by a factor of less than k. On an intuitive level, a decrease in w arising from an increase in ˜n is dampened, in part, because decreasing w increases the load per wavelength channel XC, resulting in larger total queueing delay. Decreasing w, furthermore, results in less statistical smoothing of flows arising from multiple wavelength channels working in concert to serve flows. Lastly, we note that, as ˜n grows large, each of the colored curves asymptotically approaches the value of w required for an analogous network in which each DN is connected to the MAN with 2f fibers (i.e., X = L) to achieve the same performance.

The previous numerical results and discussion provide justification for narrowing the space of design alternatives in the design of OFS networks. First, equipping each DN with 2f fibers for bidirectional communication provides little performance benefit for large MANs in which there are hundreds of DNs per MAN. This observation, coupled with the aforementioned practical concerns regarding network upgrades, render this design alternative less attractive than the case of two fibers per DN. Wavelength conversion was found to provide a moderate performance, with the benefit decreasing for large MANs. In spite of this performance benefit, use of this technology in individual DNs may be imprudent, as the present-day costs of the relevant technologies are prohibitive.

VII. CONCLUSION

In this work, we began by proposing a simple scheduling algorithm to arbitrate access to network resources for inter-MAN OFS communication. Using this algorithm, we then conducted an approximate throughput-delay analysis of OFS networks, and explored the tradeoffs in the OFS architecture design space.

The work in this paper may be extended in various directions. First, generalizing the number of wavelength channels in MAN links to be less than that required for a dedicated channel to exist for every inter-MAN OFS channel is of practical importance. Secondly, in future networks, traffic with different quality of service (QoS) requirements will likely exist. Incorporating a class of best-effort traffic with a long-term throughput requirement, but no delay requirement, would be a straightforward extension in that this traffic would appear “invisible” to higher-priority traffic. Lastly, whereas this work focused on a performance analysis of inter-MAN OFS communication, future work should address the case of intra-MAN communication.

REFERENCES