On the throughput-cost tradeoff of multi-tiered optical network architectures

Guy Weichenberg[†⋆], Vincent W. S. Chan[†], and Muriel Médard[†]
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology

***Abstract*— In this work, we conduct a throughput-cost study of several optical network architectures: Optical Flow Switching (OFS), Tell-and-Go (TaG), Electronic Packet Switching (EPS), and Generalized MultiProtocol Label Switching (GMPLS). The simple, multi-tiered optical network that we consider comprises two groups of users, each in a distinct metropolitan-area network (MAN), which wish to communicate over a wide-area network (WAN). Our network cost model focuses on initial capital expenditure: transceiver, switching, routing, and amplification costs. Our results indicate that: OFS is the most scalable architecture of all, in that it is most cost-efficient when the average user data rate is high and the number of users in the network is large; EPS is most sensible when the product of the number of users and the average user data rate is low; the GMPLS architecture, which is conceptually intermediate to EPS and OFS, is optimal when the product of the number of users and the average user data rate is moderate; and, finally, there does not exist an optimal regime for TaG.**

I. INTRODUCTION

Optical networking first emerged as copper and microwave radio links in communication networks were replaced by optical fibers. This decision, motivated by the enormous information capacity of optical fiber, prevented transmission links from acting as information bottlenecks in communication networks. Network nodes, which operated purely in the electronic domain, thus became the point of congestion in communication networks when internet bandwidth demand exploded. To make matters worse, the heterogeneity of these networks required costly optical-electronic-optical (OEO) conversions at nodes. While recent years have witnessed the development of many novel optical networking devices, electronics has remained the clear choice with which to carry out logical operations at network nodes. Thus, any sensible architecture for an expansive, terrestrial communication network in the near future will necessarily incorporate electronic and optical technologies. Network designers are therefore faced with the task of judiciously (i.e., economically) integrating optical and electronic technology into a capable network architecture. In particular, the objective should be to design a network with excellent scalability: a decreasing cost per user, per unit of traffic, as the number of users and individual user bandwidth demand increase.

Our aim in this work is to address this very question of how to best use traditional electronic and emergent optical technology to create a scalable, expansive, terrestrial network. Specifically, we conduct a throughput-cost comparison of several prominent network architectures that incorporate varying degrees of electronics and optics: Optical Flow Switching (OFS), Tell-and-Go (TaG), Electronic Packet Switching (EPS), and Generalized MultiProtocol Label Switching (GMPLS). The context in which we compare these different network architectures is a simple network comprising two large groups of users, located in different MANs, which wish to communicate over a WAN. Our network model, though simple in that it only considers the communication of two sets of users across a WAN, is a building block for more complex network topologies, and, more importantly, captures the essence of the throughput-cost tradeoffs of these more complex networks.

One limitation of our study is that it focuses on communication across the WAN, and therefore provides limited insight into how networks should be optimally designed to accommodate intra-MAN and intra-local-area network (LAN) traffic. Furthermore, we recognize that, while the throughput metric is important, it is not the only performance criterion by which a network should be assessed. Delay, for example, is another key performance metric that is not addressed in this work. Finally, our cost model focuses on the initial capital expenditure of a network, and neglects the ongoing operational costs, which may constitute a significant portion of a network's cost and may vary substantially from architecture to architecture.

This work is organized as follows. In the next section, we outline the general structure of our network and cost model. In Section III, we discuss the candidate network architectures and present their throughput-cost characterizations. We conduct a throughput-cost comparison of the different architectures with realistic network parameters in Section IV. We conclude the work in Section V.

II. NETWORK AND COST MODEL

The simple, multi-tiered optical network that we consider in this work comprises two MANs, each with n_m/η_u users ($0 < \eta_u \leq 1$)—of which n_m are active at any time—which wish to communicate over a WAN. Users are grouped into LANs, or distribution networks, which, in the cases of the OFS and TaG architectures, are passive, optical broadcast networks. The number of users per distribution network in these two architectures will be limited to a few hundred owing to minimum power requirements for optical detection at end users. For the MAN, we assume that the physical topology connecting its m nodes is arbitrary, but identical for each architecture. The MAN node design (i.e., use of an optical cross-connect (OXC) or a router), however, is dependent upon the network architecture. Although our network model

---

The research in this paper was supported by: [†]National Science Foundation under Grant No. ANI-0335217 and DARPA under Award No. MDA972-02-2-0021; and [⋆]Natural Sciences and Engineering Research Council of Canada.

Notation	Definition
c_{amp}	EDFA cost per wavelength per WAN hop
c_{mux}	Passive (de)multiplexing cost for each end user transceiver
c_{sch}	MAN scheduling processor cost
C_x	Total cost of network x, $x \in \{ofs, tag, eps, gmpls\}$
η_u	Ratio of active users to total users in a MAN
γ_{agg}	Port utilization in a LAN aggregator switch
$\gamma_{r,m}$	Port utilization in a MAN router
$\gamma_{r,w}$	Port utilization in a WAN router
h_m	Avg. no. of hops from a MAN node to closest WAN node
h_w	Avg. no. of hops between WAN nodes
m	Number of nodes in a MAN
n_l	Number of active users in a LAN distribution network
n_m	Number of active users in a MAN
p_{agg}	Bidirectional LAN aggregator switch port cost
p_{eth}	Bidirectional Ethernet switch port cost
p_{oxc}	Bidirectional OXC port cost
$p_{r,m}$	Bidirectional MAN router port cost
$p_{r,w}$	Bidirectional WAN router port cost
r	Wavelength channel data rate
S_u	Avg. active user data rate
$t_{n,l}$	Optical nontunable, long-haul transceiver cost
$t_{n,s}$	Optical nontunable, short-haul transceiver cost
t_e	Electronic transceiver cost
t_s	$t_{n,s}$ or t_e, depending on the channel rate and the architecture
$t_{t,l}$	Optical tunable, long-haul transceiver cost
w	$n_m S_u / r$, total wavelength resources consumed by active users
w_x	No. of channels in network x, $x \in \{ofs, tag, eps, gmpls\}$

TABLE I

TABLE OF NOTATION USED

is restricted to have just two MANs, this simple model may represent a subset of a more complex network. Thus, the analysis of even such a simple case may serve as a useful building block for more complex network analyses.

For the n_m users assumed to be active at any given time in each MAN, we will assume a uniform traffic demand: each active user wishes to communicate with each active user in the other MAN at the same data rate. We will let S_u, our throughput metric in this work, denote the aggregate average data rate at which each active user transmits or receives. An end user is equipped with multiple transceivers, each of which can transmit at a peak rate of r, which is the wavelength channel rate in the network. When buffers exist in the network, whether at end users or within switches or routers, they are infinite; that is, data loss due to buffer overflow never occurs. In addition, we allow for switches, routers, and OXCs within the LAN, MAN, and WAN to operate below their full line capacities (parameterized by γ), which is consistent with present-day operation of these devices[1].

In our cost model, we address the transceiver, switching, scheduling, routing, and optical amplification costs of the network entailed by bidirectional communication among the MAN users. In accounting for the switch and router costs, we fix a reasonable operating line rate (e.g., 10 Gbps) and device size (e.g., 16×16 ports), and then assume that the device cost is a linear function of port count. This assumption permits us to employ per port costs for switches and routers in our cost model, although the accuracy of the resulting analysis diminishes when large perturbations from this operating point exist. With respect to optical amplification costs within the WAN, we assume that an erbium-doped fiber amplifier (EDFA) is required approximately every 50 km, and that a single EDFA is capable of amplifying 100 wavelength channels simultaneously. We omit fiber plant costs—digging, cabling, leasing, and right-of-way costs—as we assume that they are approximately the same for all of the architectures considered. Furthermore, we omit operational costs of the network, which we recognize to be an important cost component.

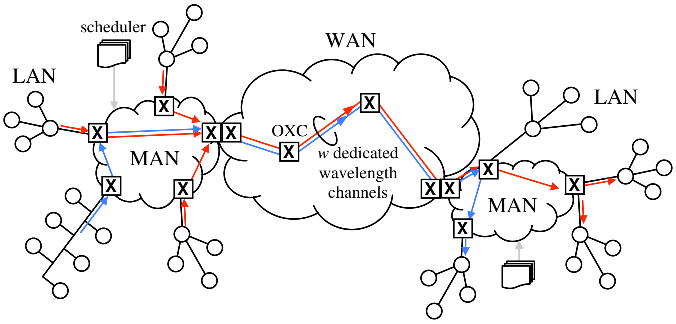


Fig. 1. Illustration of the OFS implementation.

Most of the notation used in the remainder of this work is compiled in Table I.

III. THE CANDIDATE NETWORK ARCHITECTURES

In this section, we operationally describe the four network architectures considered in this work and characterize their throughput-cost tradeoffs.

A. *OFS*

In the OFS network architecture, transmission of data is carried out in a scheduled manner between end users, akin to circuit switching, albeit for shorter durations [1], [2]. In order to schedule data transmission across the WAN, users communicate, via an off-band control plane, with the scheduling processor assigned to their respective MANs. These two scheduling processors, in turn, coordinate transmission of data across the WAN in an off-band control plane.

Motivated by the minimization of network management and switch complexity in the network core, we require that flows be serviced as indivisible entities and that WAN bandwidth be dedicated for each MAN pair. The latter assumption of dedicated bandwidth, however, is not a a general feature of OFS networks [1], [2]. Indeed, OFS is considered a centralized transport architecture in that coordination is required for logical topology reconfiguration, although we anticipate that such reconfiguration would occur on coarse time scales. This is a justification for our assumption that WAN bandwidth is statically dedicated for each MAN pair.

We further assume that wavelength conversion is not used in the network, although this too is not a requirement of the OFS architecture. 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. Note that, in OFS networks, unlike packet switched networks, all queueing of data occurs at the end users, thereby obviating the need for buffering in the network core.

---

[1] Sub-capacity operation of these devices ensures reasonable delay of data transactions, which is a consideration that is beyond the scope of this work. Routers and switches, furthermore, operate well below their capacities because of the computational bottleneck imposed by the network processing unit.

The OFS implementation of the simple, two MAN network considered in this work is illustrated in Figure 1. Each user is equipped with at least one tunable, long-haul transceiver for data communication and one short-haul transceiver with which to communicate with the scheduler[2]. In the MAN and WAN, switching of data occurs all-optically via OXCs.

Throughput-cost characteristics: In our previous work [3, Theorem 2], we characterized the achievable throughput in OFS networks by viewing such networks as generalized switches. Using these results, it can be shown that, under uniform traffic, each active user can achieve an average data rate of $w_{ofs}r/n_m$, where w_{ofs} wavelength channels are dedicated to the MAN pair in the WAN. Conversely, since WAN bandwidth is allocated at wavelength granularity, $w_{ofs} = \lceil n_m S_u/r \rceil$ dedicated wavelength channels are required in the WAN to accommodate an average active user data rate of S_u. The average wavelength channel utilization in the WAN is therefore $n_m S_u/(r w_{ofs})$.

The throughput-cost relationship of the OFS implementation of our simple network is given by:

$$C_{ofs} = \frac{2n_m}{\eta_u}\left\lceil \frac{S_u}{r} \right\rceil t_{t,l} + \left(\frac{2n_m}{\eta_u} + 2\right) t_s + 2c_{sch} + \frac{2n_m}{\eta_u}\left\lceil \frac{S_u}{r} \right\rceil c_{mux} + 2w_{ofs} m \left(h_m + 1\right) p_{oxc} + w_{ofs}\left(h_w + 1\right) p_{oxc} + w_{ofs} h_w c_{amp},$$

where n_l is the number of active users in a distribution network. In this expression, the first term represents the cost of the $\lceil S_u/r \rceil$ transceivers at the $2n_m/\eta_u$ users in the network; the second and third terms represent the cost of scheduling in the two MANs; the fourth term represents the cost of broadcasting/combining data in the distribution networks; the fifth and sixth terms represent the switching cost in the MANs and WAN, respectively; and the seventh term represents the cost of amplification in the WAN.

Owing to the fact that, in our OFS model, resources are dedicated for MANs communicating across a WAN, the throughput-cost characteristic of each MAN pair derived above would approximately hold in a more complex network containing many MANs.

B. TaG

The TaG architecture, like the OFS architecture, is one in which data is communicated between end users without being buffered or processed at intermediate nodes [4], [5]. The major difference between TaG and OFS, however, is that transmission of data is carried out in a random-access fashion in TaG rather than in a scheduled fashion. This architecture, while simpler owing to the absence of scheduling and coordination, allows for collisions of data to occur both on the wavelength channels and at the receiving users. The TaG implementation of the two MAN network is depicted in Figure 2. As in OFS, we assume an arbitrary OXC-connected MAN topology and dedicated wavelength channels in the WAN. We remark that, had we allowed for random-access to WAN wavelength resources instead of dedicated resources, then the WAN network architecture would be that of optical burst switching (OBS).

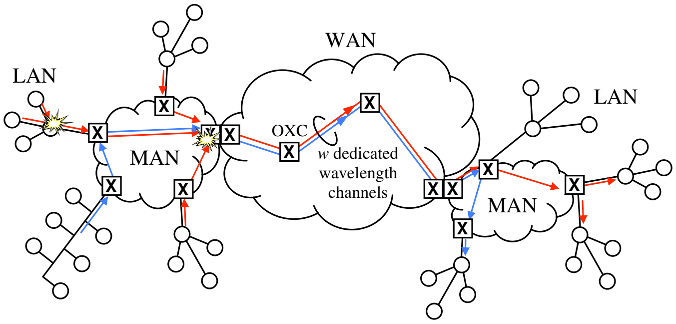


Fig. 2. Illustration of the TaG implementation.

Throughput-cost characteristics: We view each wavelength channel in each distribution network as a multiple-access system. The w_{tag} wavelength channels are coupled in that, at any instant in time, a transmitting user may only be transmitting on one wavelength channel with each transceiver. For the sake of analytic tractability, however, we neglect this coupling constraint. Within each distribution network, we therefore treat each wavelength channel as an independent (possibly slotted) Aloha network with variable length bursts. For the special case of constant length bursts[3] in which the burst length is a large multiple of the slot length (if time is slotted), the throughput per wavelength channel emerging from a distribution network is approximately that of unslotted Aloha: $S_{d,1} = Ge^{-2G}$, where G is the aggregate transmission—fresh arrival and retransmission—intensity.

The manner in which traffic is shaped by the MAN network after emerging from the distribution networks is dependent upon the topology of the MAN. However, we provide an approximate analysis that is not topology dependent. In particular, we neglect propagation delay, and we assume that the logical topology is that of a tree in which OXC ports mediate access to the MAN wavelength channels. This latter assumption of mediation of the channel by OXC ports enables channel capture; that is, if a wavelength channel is free when a burst transmission is attempted on it, then the channel is reserved for the duration of the burst. We assume that each distribution network produces, independent of other distribution networks, a sequence of independent transmit and idle states, in which the transmit state represents both the useful and garbled data (owing to collisions) from its constituent end users. The throughput on each wavelength channel emerging from the MAN is given by:

$$S_{m,1} = \frac{n_m}{n_l} S_{d,1} \left(e^{-G}\right)^{\frac{n_m}{n_l} - 1},$$

where we have assumed independence among the $\frac{n_m}{n_l}$ distribution networks in the MAN.

In our network, transactions on the w_{tag} WAN wavelength channels are coupled in that they may contend for the same transceivers in the receiving MAN. We make the approximation

---

[2] Alternatively, a single tunable, long-haul transceiver can perform both data communication and scheduling.

[3] For a fixed mean burst length, constant length bursts can be shown to maximize throughput [6].

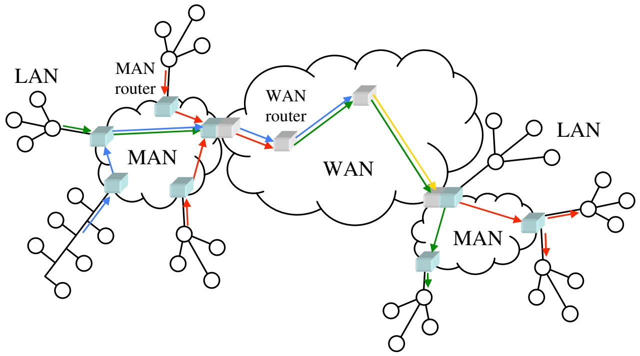


Fig. 3. Illustration of the EPS implementation.

that, from the perspective of each of these transceivers, the data on each of the w_{tag} wavelength channels generates a sequence of independent transmit and exponentially distributed idle states, with expected transmit length $\overline{L}$ (corresponding to *useful* data) and expected idle length $\overline{L}\,[n_m/S_{m,1}-1]$. Following previous work [7], it can be shown that, for the case of constant length bursts, the fraction of time that each WAN wavelength channel contains useful data (i.e., data not involved in a receiver collision) is:

$$S_{tag} = \frac{n_m \rho}{(1+\rho)^{w_{tag}}} e^{(1-w_{tag})\rho},$$

where $\rho = [n_m/S_{m,1}-1]^{-1}$. This yields a user throughput of $S_u = w_{tag}S_{tag}/n_m$. We remark that, if there existed other MANs in the network, then the achievable user throughput would be further diminished by receiver collisions involving data from users in other transmitting MANs. Thus, the above analysis serves as a "best case" throughput bound for more complex networks with multiple MANs.

The cost-throughput tradeoff of the TaG configuration is similar to that of OFS except for: i) the absence of scheduling equipment cost, and ii) the number of transceivers is linked to the *attempted* user intensity instead of the user throughput:

$$C_{tag} = \frac{2n_m}{\eta_u}\left\lceil \frac{w_{tag}G}{n_l} \right\rceil t_{t,l} + \frac{2n_m}{\eta_u}\left\lceil \frac{w_{tag}G}{n_l} \right\rceil c_{mux} + w_{tag}h_w c_{amp} + 2w_{tag}m\,(h_m+1)\,p_{oxc} + w_{tag}\,(h_w+1)\,p_{oxc}.$$

Note that w_{tag}, which is a function of S_u, provides the link between throughput and cost for the TaG architecture.

C. EPS

In the EPS implementation of the two MAN network, drawn in Figure 3, the LAN and MAN architectures resemble that of present-day Ethernet networks. Specifically, data from an end user is first aggregated at a LAN Ethernet switch and is then further aggregated at a LAN aggregator switch which multiplexes data from several LANs. Traffic is then fed into the closest MAN router, and subsequently routed to the MAN's WAN router. Within the arbitrarily connected MAN, electronic transceivers are used whenever possible (e.g., when line rates are sub-Gbps), as they are far less expensive than optical transceivers. As in the MAN, data in the WAN is statistically multiplexed anew at each hop en route to the destination MAN. In the WAN, data is always sufficiently aggregated and intense to warrant optical transmission. Processing of data, however, is carried out at *each* hop in the WAN in the electronic domain.

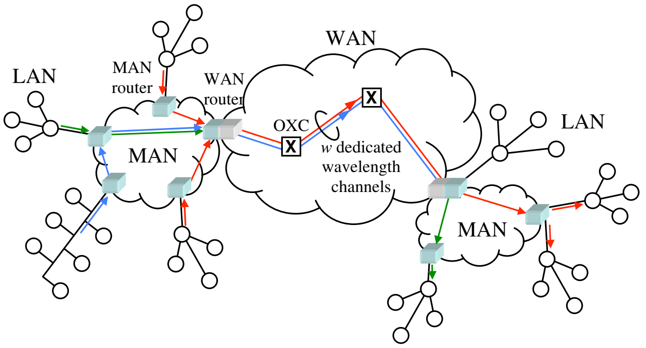


Fig. 4. Illustration of the GMPLS implementation.

Throughput-cost characteristics: In [3], [8], it is shown that statistical multiplexing of data in packet switched networks enables 100% utilization of communication channels. That is, unlike OFS and TaG which employ dedicated resources and have no buffering in the WAN and may thus waste wavelength channel capacity, statistical multiplexing in EPS allows for wavelength channels to be utilized to their capacity limits, or to their scaled limits dictated by router computational bottlenecks. Thus, if we assume that wavelength channels are loaded to their γ-scaled capacity limits, then the cost of supporting a certain amount traffic is exactly equal to a $1/\gamma$ scaling of the cost of resources that are consumed in serving this traffic. To be precise, within the MAN, the assigned amount of wavelength resources consumed is $w/\gamma_{r,m}$ where $w = n_m S_u/r$; and within the WAN, it is $w_{eps} = w/\gamma_{r,w}$. The cost-throughput relationship of the EPS configuration is therefore given by:

$$C_{eps} = \frac{2n_m}{\eta_u}\left\lceil \frac{S_u}{r} \right\rceil t_s + \frac{2n_m}{\eta_u}\left\lceil \frac{S_u}{r} \right\rceil (p_{eth}+t_s) + 2\left\lceil \frac{n_l S_u}{r\gamma_{agg}} \right\rceil \frac{n_m}{n_l}(p_{agg}+t_s) + \frac{2w}{\gamma_{r,m}}(h_m+1)(p_{r,m}+t_s) + w_{eps}(h_w+1)(p_{r,w}+t_{n,l}) + w_{eps}h_w c_{amp}$$

where $t_s = t_{n,s}$ or t_e, depending on the wavelength channel data rate. In this expression, the first term represents the cost of the $\lceil S_u/r \rceil$ transceivers at the $2n_m/\eta_u$ users in the network; the second and third terms represent the cost of multiplexing (demultiplexing) data from (to) the end users in the LANs; the fourth and fifth terms represent the cost of routing data in the MANs and WAN, respectively; and the sixth term represents the cost of amplification in the WAN.

As discussed previously, statistical multiplexing of data in EPS networks allows us to assign the cost of communication of the MAN pair as the $1/\gamma$-scaled cost of network resources consumed, independent of the other users that may be sharing these resources. Hence, the throughput-cost characteristic of each MAN pair derived above would approximately hold in a more complex network containing many MANs.

D. GMPLS

The version of GMPLS considered here is conceptually intermediate to OFS and EPS. Specifically, the LAN and MAN

$c_{amp} = \$2000$	$c_{mux} = \$30$	$c_{sch} = \$5000$
$\eta_u = 0.1$	$\gamma_{agg} = 0.8$	$\gamma_{r,m} = \gamma_{r,w} = 0.3$
$h_m = 3$ hops	$h_w = 5$ hops	$m = 15$ nodes
$n_l = 30$ users	$n_m = 10,000$ users	$p_{agg} = \$300$
$p_{eth} = \$100$	$p_{oxc} = \$10,000$	$p_{r,m} = \$60,000$
$p_{r,w} = \$125,000$	$r = 10$ Gbps	$t_e = \$30$
$t_{n,l} = \$1000$	$t_{n,s} = \$300$	$t_{t,l} = \$2000$

TABLE II

TABLE OF PARAMETERS USED IN SECTION IV.

in GMPLS is identical to that of EPS, while the WAN design is similar to OFS in that all-optical transmission along dedicated wavelength channels passing through OXCs is employed in order to circumvent electronic processing at intermediate nodes[4]. At the interface of the MAN and the WAN exist ingress and egress routers that are responsible for assembling and disassembling large blocks of data, respectively. The GMPLS architecture is illustrated in Figure 4 for the two MAN network.

Throughput-cost characteristics: In the MAN, traffic is treated in the same way as in EPS, so the amount of wavelength resources assigned is $w/\gamma_{r,m}$ where $w = n_m S_u / r$. However, data traverses the WAN through a fully optical path, akin to OFS. Thus, data is routed through the WAN using $w_{gmpls} = \lceil w/\gamma_{r,w} \rceil$ dedicated wavelength channels. The cost of the GMPLS configuration is therefore given by:

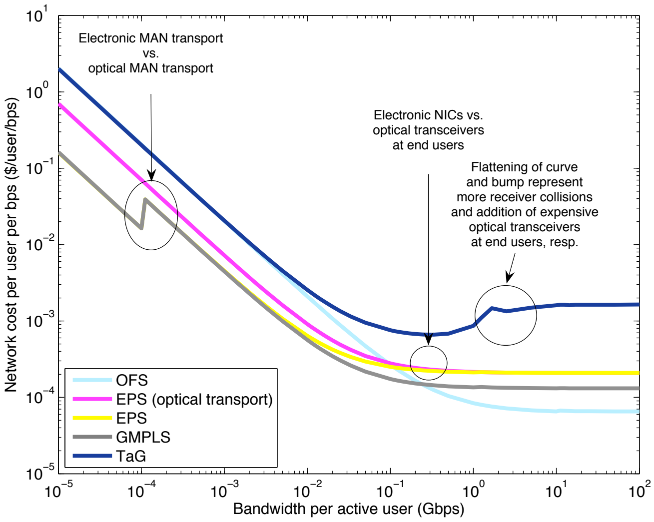
$$\begin{aligned} C_{gmpls} = {} & \frac{2n_m}{\eta_u} \left\lceil \frac{S_u}{r} \right\rceil t_s + \frac{2n_m}{\eta_u} \left\lceil \frac{S_u}{r} \right\rceil (p_{eth} + t_s) \\ & + 2 \left\lceil \frac{n_l S_u}{r \gamma_{agg}} \right\rceil \frac{n_m}{n_l} (p_{agg} + t_s) + \frac{2w}{\gamma_{r,m}} (h_m + 1)(p_{r,m} + t_s) \\ & + 2 w_{gmpls} (p_{r,w} + t_{n,l}) + w_{gmpls} (h_w - 1) p_{oxc} \\ & + w_{gmpls} h_w c_{amp}. \end{aligned}$$

By the same reasoning as for OFS and EPS, the above throughput-cost tradeoff of each MAN pair would approximately hold in more complex networks with many MANs.

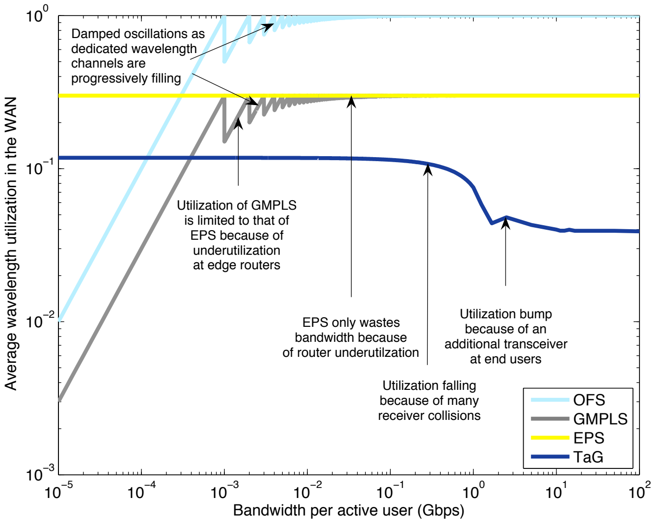
IV. ARCHITECTURAL THROUGHPUT-COST COMPARISON

In this section, we carry out a throughput-cost comparison of the four network architectures. The cost and architectural parameters used, which reflect the state of present-day networks (e.g., [9]), are summarized in Table II.

Figure 5 depicts some results of our throughput-cost study. In Figure 5(a), we plot network cost per user per bps versus average active user data rate. The plot indicates that when users have data rates less than approximately 200 Mbps, the EPS and GMPLS architectures have the lowest normalized cost. Intuitively, this is because relatively little expensive electronic equipment is necessary to support the aggregate traffic in these architectures; whereas in OFS and TaG, each end user is equipped with an expensive long-haul optical transceiver, even at low data rates. Beyond data rates of 200 Mbps, however, we see that OFS is the most cost-efficient architecture because the high cost of transceivers at end users is offset by the low cost of provisioning wavelengths in the WAN. In the EPS and GMPLS architectures, by contrast, high data rates require much expensive electronic equipment in the MAN and WAN.



(a)



(b)

Fig. 5. The operating assumptions given in Table II are made. The "EPS (optical transport)" architecture is identical to the EPS architecture, except that transmission is *always* carried out in optics, even at low data rates.

In Figure 5(b), we plot average WAN wavelength channel utilization versus average active user data rate. We note that, as the data rate increases, the average wavelength utilization for OFS converges to unity, whereas the other architectures' utilizations are clearly bounded away from unity. In the cases of EPS and GMPLS, utilization is bounded away from unity solely because of the fact that routers are underutilized in these schemes. In the case of TaG, the underutilization of the WAN wavelength channels is intrinsic to the architecture: random-access to the network's resources results in channel and receiver collisions which waste channel capacity.

In Figure 6, we indicate the cost-optimal network architecture as a function of the number of users per MAN and the average active user data rate. We observe, for the same reasons discussed in reference to Figure 5(a), that EPS and GMPLS dominate for low to moderate data rates. However, this figure further indicates that EPS performs best when the product of the number of users and average active user data rate is low. Thus, EPS is most sensible when there are many users at very

---

[4] GMPLS, in its most general form, permits switching in other domains besides wavelength/frequency.

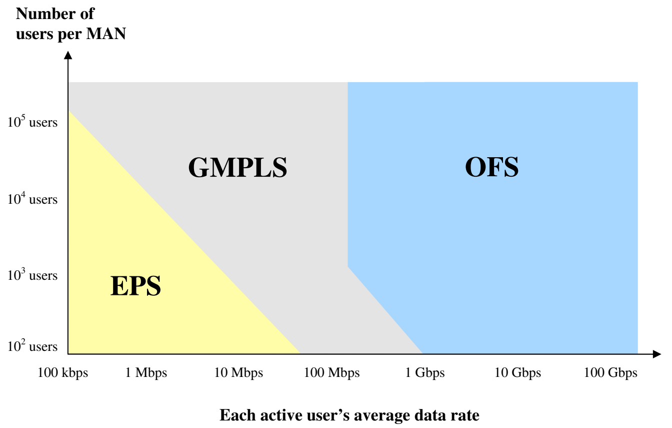


Fig. 6. Cost-optimal architecture as a function of average active user data rate and number of users. We use the operating assumptions in Table II.

low data rates (as in the case of today's Internet), very few users at moderate data rates, or anywhere in between these two extremes. GMPLS performs best when the product of the number of users and average active user data rate is moderate. These observations are intuitive since aggregate traffic is light in the former case, in which case it is wasteful to provision entire wavelengths in the WAN; whereas in the latter case, aggregate traffic is heavy, in which case provisioning entire wavelengths is sensible. In fact, it can be shown that, under heavy aggregate traffic, the cost difference between EPS and GMPLS scales proportionally to the product of the aggregate traffic and the difference in cost between a router port (with a transceiver termination) and an OXC port.

Another immediate observation is that, at high data rates, regardless of the number of users, OFS always dominates, implying that *OFS is the most scalable architecture of all.* In the high user data rate regime, aggregate traffic is always high, so requiring electronic equipment to support this traffic in the network—even if only in the MAN—is expensive. In fact, in can be shown that the switching cost difference in the MAN and WAN is approximately proportional to the product of the aggregate traffic and the difference in cost between a router port (with a transceiver termination) and an OXC port. This trend is furthered by the higher LAN switching costs in the EPS and GMPLS architectures. This trend, however, is offset somewhat by the higher transceiver cost at end users in the OFS architecture. In particular, the OFS implementation requires a tunable long-haul transceiver and possibly a separate short-haul transceiver for scheduling at each end user, whereas EPS and GMPLS only require a nontunable short-haul transceiver at each end user. The overall effect, as shown in Figure 6, is a higher cost under EPS and GMPLS than under OFS as the number of users and the average active user data rate increase.

We note that there does not exist a regime of optimality for TaG. This is expected since the low cost of scheduling in OFS yields great performance benefit relative to the otherwise identical TaG architecture. More specifically, the savings in scheduling equipment in TaG is dwarfed by the cost of supporting many more wavelength channels than in OFS. Although delay is beyond the scope of this work, we remark that there exists a delay-utilization-blocking probability tradeoff for OFS (with finite buffers) and TaG. We expect this tradeoff to be less attractive for TaG since delay is expected to be especially poor in moderate/heavy traffic owing to many retransmissions. Segmentation and "tree" algorithms may improve TaG performance, but this comes at the expense of protocol complexity.

Figure 6 also suggests that a hybrid architecture may be the most sensible design of all, especially when user demands are heterogeneous. For example, a network in which high data rate users employ OFS and low data rate users employ EPS or GMPLS would perform better than a network in which only one of the aforementioned architectures is employed.

As a final remark, note that, while the precise boundary positions in Figures 5 and 6 are sensitive to the exact parameter values in Table II, the general trends observed are manifestations of the present-day cost structures of the architectures and their device building blocks. Thus, in the absence of disruptive technologies with radically different cost structures, we expect the trends observed in these figures to hold for a wide range of parameters and for more complex networks.

V. CONCLUSION

In this work, we studied the throughput-cost characteristics of several optical network architectures—OFS, TaG, EPS, and GMPLS—in the context of a simple, multi-tiered optical network. Given the present-day cost structures of the architectures and their device building blocks, we show that OFS is the most scalable architecture of all, in that it is most cost-efficient when the average user data rate is high and the number of users in the network is large; EPS is most sensible when the product of the number of users and the average user data rate is low (as in the case of today's Internet); the GMPLS architecture, which is conceptually intermediate to EPS and OFS, is optimal when the product of the number of users and the average user data rate is moderate. Our work, while not exhaustive in its consideration of network architectures and topologies, is valuable nevertheless in that it provides insight into the key properties of networks that impact their scalability.

REFERENCES

[1] V. W. S. Chan, K. L. Hall, E. Modiano, and K. A. Rauschenbach, "Architectures and technologies for high-speed optical data networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 16, no. 12, pp. 2146–2168, Dec. 1998.

[2] V. W. S. Chan, G. Weichenberg, and M. Médard, "Optical flow switching," *Proceedings of the Workshop on Optical Burst Switching*, Oct. 2006.

[3] G. Weichenberg, V. W. S. Chan, and M. Médard, "On the capacity of optical networks: A framework for comparing different transport architectures," *Proceedings of IEEE Infocom*, 2006.

[4] I. Widjaja, "Performance analysis of burst admission-control protocols," *IEE Proceedings Communications*, vol. 142, no. 1, pp. 7–14, Feb. 1995.

[5] J. Turner, "Terabit burst switching," *Journal of High Speed Networks*, vol. 8, no. 1, pp. 3–16, Jan. 1999.

[6] S. Bellini and F. Borgonovo, "On the throughput of an ALOHA channel with variable length packets," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1932–1935, Nov. 1980.

[7] M. J. Ferguson, "An approximate analysis of delay for fixed and variable length packets in an unslotted Aloha channel," *IEEE Transactions on Communications*, vol. 25, no. 7, pp. 644–654, July 1977.

[8] M. Marsan, P. Giaccone, E. Leonardi, and F. Neri, "On the stability of local scheduling policies in networks of packet switches with input queues," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 4, pp. 642–655, May 2003.

[9] S. Sengupta, V. Kumar, and D. Saha, "Switched optical backbone for cost-effective scalable core IP networks," *IEEE Communications Magazine*, vol. 41, no. 6, pp. 60–70, June 2003.

EDFA	erbium-doped fiber amplifier
EPS	Electronic Packet Switching
GMPLS	Generalized MultiProtocol Label Switching
LAN	local-area network
MAN	metropolitan-area network
OBS	optical burst switching
OEO	optical-electronic-optical
OFS	Optical Flow Switching
OXC	optical cross-connect
TaG	Tell-and-Go
WAN	wide-area network