The Role of Optics and Electronics in High-Capacity Routers

Rodney S. Tucker, *Fellow, IEEE*

Invited Paper

***Abstract*—This paper examines the role of optical and electronic technologies in future high-capacity routers. In particular, optical and electronic technologies for use in the key router functions of buffering and switching are compared. The comparison is based on aggressive but plausible estimates of buffer and switch performance projected out to around 2020. The analysis of buffer technologies uses a new model of power dissipation in optical-delay-line buffers using optical fiber and planar waveguides, including slow-light waveguides. Using this model together with models of storage capacity in ideal and nonideal slow-light delay lines, the power dissipation and scaling characteristics of optical and electronic buffers are compared. The author concludes that planar integrated optical buffers occupy larger chip area than electronic buffers, dissipate more power than electronic buffers, and are limited in capacity to, at most, a few IP packets. Optical fiber-based buffers have lower power dissipation but are bulky. The author also concludes that electronic buffering will remain the technology of choice in future high-capacity routers. The power dissipation of high-capacity optical and electronic cross connects for a number of cross connect architectures is compared. The author shows that optical and electronic cross connects dissipate similar power and require a similar chip area. Optical technologies show a potential for inclusion in high-capacity routers, especially as the basis for arrayed-waveguide-grating-based cross connects and as components in E/O/E interconnects. A major challenge in large cross connects, both optical and electronic, will be to efficiently manage the very large number of interconnects between chips and boards. The general conclusion is that electronic technologies are likely to remain as integral components in the signal transmission path of future high-capacity routers. There does not appear to be a compelling case for replacing electronic routers with optically transparent optical packet switches.**

***Index Terms*—Buffer memories, optical delay lines, optical switches, packet switching, slow light.**

I. Introduction

THE VOLUME of traffic on the Internet continues to expand rapidly. This expansion drives a continued growth of the telecommunications infrastructure through the deployment of new transmission and switching equipment. Studies of the costs and scalability of various network architectures [1] show that as the size and capacity of the core network grows, packet switching will remain a key component of the network infrastructure. This is because, packet switching provides an end-to-end connectivity on a subwavelength basis. In addition, the packet switching provides traffic grooming at the subwavelength level, and this is a significant factor in ensuring a maximum utilization of network resources. In today's network, electronic routers provide this subwavelength connectivity and grooming, and to date, electronic routers have kept up with the demands of traffic growth. However, there is some concern in the research community that the achievable capacity of the network will eventually be limited by so-called electronic bottlenecks in electronic routers. Also of concern is the large power consumption and heat dissipation associated with large routers. The question of power consumption is already a major issue [2], [3] and is likely to become even more problematic as routers become even larger.

A number of solutions have been proposed to address the power dissipation and scaling problems in electronic routers. These solutions include replacing the electronic routers with all-optical packet switches (OPS) or photonic packet switches, in which optical packets are buffered and routed in optical form [4]–[8], or with all-optical burst switches (OBS), in which there is no buffering [9]. It is not yet clear whether OBS will be a viable alternative to electronic routing because of its relatively low throughput [10], which needs to be compensated by overbuilding the network [11].

An important feature of OPS is that it is inherently optically transparent [5]. Optical transparency offers data format independence and the opportunity to increase the data rate to values that are beyond the capabilities of electronics. Like any all-optical networking technology, optically transparent OPS will require optically transparent regenerators to circumvent the effects of accumulated optical impairments and noise. With continuous advances in the speed and functionality of electronics, it remains an open question as to whether optically transparent packet switching will offer real advantages in practical networks. A key challenge in finding a technically feasible solution to optical packet switching is the lack of an adequate optical buffering technology. However, a recent surge of new research results in the area of slow light [12]–[18] has opened up the possibility of improved optical buffer performance. While OPS appears on the surface to be appealing, it is not clear whether OPS can be scaled up to realistic sizes. This

Manuscript received May 30, 2006; revised September 10, 2006.

The author is with the ARC Special Research Centre for Ultra-Broadband Information Networks, Department of Electrical and Electronic Engineering, University of Melbourne, Melbourne 3010, Australia (e-mail: r.tucker@ee.unimelb.edu.au).

Digital Object Identifier 10.1109/JLT.2006.885774

uncertainty about scalability is heightened if we consider future networks carrying orders of magnitude of more data than the current network.

The objective of this paper is to determine whether the OPS has a potential for replacing electronic packet switching in future high-capacity networks. We approach this question by comparing the capability of key functional blocks in OPS and electronic packet switches. In particular, we compare optical buffers and high-speed optical cross connects with electronic buffers and high-speed electronic cross connects. Packet switches also contain other functional blocks, such as packet synchronizers, IP lookup processors, and switch controllers [19]. However, we focus on buffers and cross connects because these key functions are fundamentally different in optical and electronic realizations of packet switches. Most other processing functions in the packet switch will be handled by electronic processors and these functions will be similar or identical in optical and electronic packet switches. Our analysis is based on future routers and uses aggressive but plausible estimates of optical and electronic device performance, projected out to around 2020.

The basis of our comparison between optical and electronic packet switching is the power dissipation and physical size of buffers and cross connects. Previous studies of optical packet switching have focused on the technical viability of different approaches using limited-scale test-bed demonstrations, and on economic comparisons between competing technologies [1], [20], [21]. There is no doubt that a limited-scale OPS is technically feasible. However, feasibility on a limited scale is not a sufficient condition for a new technology to be adopted. Ultimately, limitations on power dissipation and footprint of full-scale systems must also be satisfied. Regardless of costs and functionality, a router that consumes and/or dissipates too much power or occupies too much floor space will not find its way into the marketplace if an alternative lower power solution exists.

For the purposes of this paper, we consider electronic buffers and cross connects using future 20-nm CMOS technologies. However, our general conclusions apply to a broader range of electronic technologies including SiGe, HBT, and BiCMOS [22]. We show that the energy per stored bit in optical delay lines increases with the size of the buffer. Slow-light optical buffers show some promise for OPS applications if the buffer capacity is small ($<$ 1 kB/port, or about 1 IPv4 packet). However, if the buffers need to have significantly larger capacity than this, fiber delay lines are the only viable option for optical buffering. In general, electronic buffer technologies offer significant advantages over the projected best possible performance of slow-light and fiber optical buffers. We show that optical switch fabrics based on tunable wavelength converters (TWCs) and arrayed waveguide gratings (AWGs) potentially offer lower power solutions than electronic cross connects. However, electronics will remain a contender for the switch fabric in large routers. Our broad conclusion is that electronic technologies are likely to remain as integral components of future high-capacity routers. There does not appear to be a compelling case for replacing electronic routers with optically transparent OPS.

To provide a concrete set of comparisons, we focus here on switch/routers with a total capacity of 100 Tb/s (10^{14} b/s) and 1 Pb/s (10^{15} b/s). The results can be readily scaled to other router capacities. Where it is feasible, we have based our projections on fundamental limitations based on device physics considerations. For example, our projections of slow-light buffers are based on new insights into the fundamental physical limitations and scaling properties of slow-light optical delay lines. In general, this gives an optimistic view of the capabilities as it ignores some nonideal behavior. However, we show that the loss in waveguide devices is a key parameter. Because there is no easily determined fundamental limitation on loss in waveguides, we base our calculations on expected best experimental results. Our estimates of future performance of CMOS devices and circuits are based on the 2005 International Technology Roadmap for Semiconductors [23], using the year 2018 as our reference point.

This paper is organized as follows. In Section II, we set out the key parameters of optical-delay-line buffers using slow light and fiber. Section III compares the properties of optical and electronic buffers, and Section IV covers the key parameters of all-optical and electronic cross connect fabrics. In Section V, we compare the power dissipation and footprint of routers using optical and electronic buffers and cross connects.

II. Optical-Delay-Line Buffers

Conventional high-capacity electronic routers employ around 250 ms of buffering on each port [24]. If a 1-Pb/s router (with 25 000 ports, each at a line rate of 40 Gb/s) has 250 ms of buffer on each port, the buffer size for each port would be 10 Gb and the total capacity of all buffers in the router would be 250 Tb. Buffers of this size can, in principle, be accommodated using an electronic random access memory (RAM), but delay-line buffers of this size are not practical. A simple calculation shows that a 250-Tb buffer using fiber delay lines would require 1 Tm of optical fiber—enough fiber to reach approximately from the sun to the planet Saturn [25]. If 100 wavelengths, each containing buffered packets were multiplexed onto each fiber, the total fiber length per router could be reduced to 10 Gm. While this is a substantial reduction in total length, it remains impractically large and the complexity switches, and multiplexers would be problematic. In essence, buffers of this size are impractical because each buffer is, in fact, an ultralong-haul link. We show in Section III below that slow-light delay lines are also impractical for use in large buffers.

If there are a large number of transmission-control-protocol (TCP) sessions running simultaneously on each wavelength, it may be possible to significantly reduce the size of buffers on routers in the network [24], [26], [27]. Based on this finding, it has been suggested that using a paced TCP, the buffer capacity could be reduced to as little as ten packets per port [28] or around 2.5 μs at 40 Gb/s [24]. This could open up new opportunities for delay-line optical buffering. However, the full impact of small buffers on large networks is not yet well understood, especially in very large networks and when individual TCP sessions are at high bit rates (e.g., 1 Gb/s or

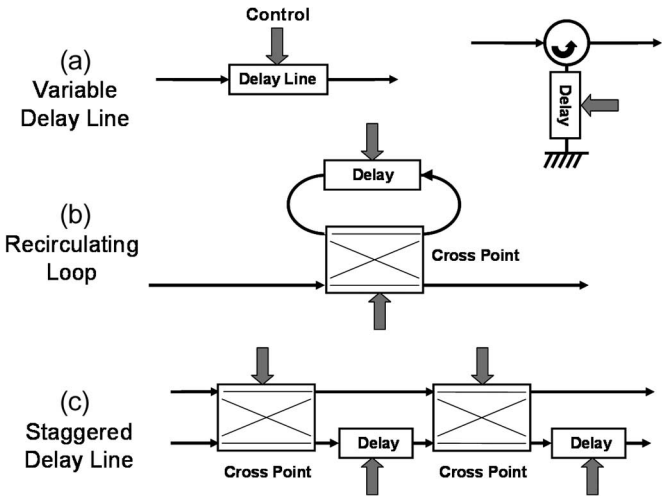

Fig. 1. Delay-line-buffer architectures. (a) Variable delay line. (b) Recirculating loop. (c) Staggered delay line.

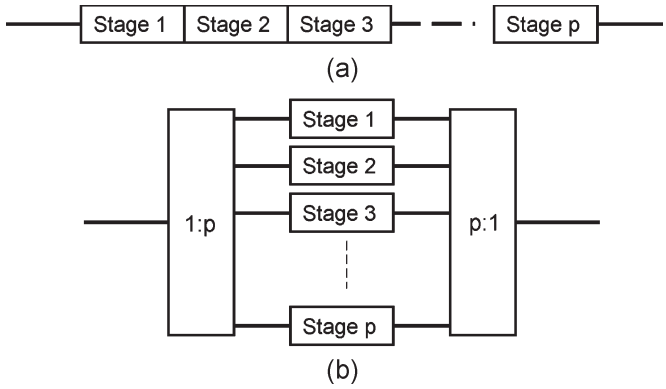

Fig. 2. (a) Serial and (b) parallel delay-line structures.

more in future fiber to the home systems). In addition, the throughput of routers with around ten packets of buffering per port is 80% or less [28] compared with close to 100% for routers with large buffers. This 80% throughput figure might seem to be quite respectable. However, in order to provide the same overall network performance as a network using routers with large buffers, networks using routers with small buffers will require the routers and the links between routers to be overbuilt to compensate for the 20% loss in throughput. This situation is similar to the situation likely to confront designers of optical burst switched networks, where an overbuilding is also necessary, and this overbuilding seriously cuts into the economically viability of OBS [11]. Clearly, more work needs to be done on resolving the question of the optimum buffer size in OPS.

Despite the myriad recent advances in optics and photonics, the search for a compact high-speed digital optical memory continues to frustrate researchers. The best available approach to buffering in optical packet switching uses optical delay lines. Fig. 1 shows four basic optical-delay-line building-block stages. Fig. 1(a) is the simplest of all delay-line buffers—a line with a variable delay. In Fig. 1, the broad arrows represent a control signal that is used to adjust the delay. It is important to note that a simple fixed delay line is not a buffer, and it cannot be used to store optical data. Buffering or storage requires that there should be some control over when the optical data are readout from the delay line [18]. Fig. 1(b) shows a variable (or fixed) optical delay line, combined with a crosspoint, in a feedback (recirculating loop) configuration. Unlike Fig. 1(a), the delay line in Fig. 1(b) does not need to be controllable because data can be readout using the controllable cross connect. Fig. 1(c) shows a feed-forward arrangement with crosspoints and delay lines. The capacity of all the structures in Fig. 1 can, in principle, be expanded by a wavelength-division multiplexing.

Buffers can be constructed by combining these (and other) building blocks either in cascade, as shown in Fig. 2(a), or in parallel as shown in Fig. 2(b). If the cascaded stages in Fig. 2(a) are simple variable delay lines of the type shown in Fig. 1(a), the cascade is a first-in-first-out (FIFO) buffer [18], in which the order of packets emerging from the buffer is the same as the order of packets entering it. The recirculating loop buffer in Fig. 1(b) can operate as a FIFO buffer, but if more than one packet is stored in the delay, the order of the packets can be changed. There is a common misconception that less waveguide delay is required in recirculating buffers than other buffer memories. However, the reality is that while a packet is circulating in the loop, no other packets can enter it. Therefore, if it is necessary to buffer a stream of closely spaced incoming packets, multiple recirculating loops are required. This can be achieved using the cascade configuration of Fig. 2(a) or the parallel configuration in Fig. 2(b).

A. Fiber Delay Lines

The most commonly used optical buffers in OPS system experiments in the literature are based on delay lines using optical fiber as the waveguide. Many researchers have investigated the application of fiber-delay-line buffers to OPS, and a wide variety of buffers based on fiber delays have been reported [4], [29]–[33]. Because the delay of a fiber delay line cannot be directly controlled, fiber-delay-line buffers require additional control components such as the crosspoint switches shown in Fig. 1. For example, a buffer that can store up to N packets will require at least N crosspoints and N fiber delay lines, each with a delay time at least equal to one packet length [see, for example, Fig. 1(c)]. In practical fiber-delay-line buffers, attenuation in the buffer is dominated by the attenuation in the crosspoints. We will examine the implications of this in Section II-C. A possible alternative to fiber delay lines is low-loss integrated silica delay lines [34]. While these waveguides are more compact than fiber delay lines, their data storage capacity is limited by waveguide losses.

B. Slow-Light Delay Lines

A recent explosion of research activity in slow-light technologies [12]–[18] has opened up new optimism for the future of delay-line optical buffers. Because of their lower group velocity, the size of the individual bits of data is smaller and the required length of the waveguide for a given storage capacity is reduced [18], [35]. While slow-light delay lines

show great promise, it is important to recognize that their capacity is limited by the delay-bandwidth product or delay-throughput product, which is determined by the particular slow-light technology [18]. The two most promising slow-light technologies for providing tunable delays in OPS buffer applications are electromagnetically induced transparency (EIT) in semiconductor quantum dots (QDs) [14] and coupled-resonator waveguides (CRWs) [36]. A third possibility is photonic crystal (PC) waveguides operated close to the photonic band edge [16]. However, this class of delay line has very large second-order dispersion that severely limits the useable bandwidth. In addition, the amplitude response shows a strong tilt across the signal passband. While there have been some successful attempts to reduce the dispersion in PC waveguides operated close to the photonic band edge [37], CRWs offer the potential of superior dispersion-limited bandwidth performance.

There are two distinct classes of slow-light delay lines—Class A and Class B [18] (Class B is also referred to as "adiabatically tunable" [38]). In Class A delay lines, the group velocity of the light is slowed across an interface between two waveguides with different group velocities. In Class B delay lines, the velocity of the light is reduced adiabatically and uniformly across a data packet [18], [38]. There is a common misconception in the literature that Class B delay lines circumvent the delay-bandwidth product limitation encountered in Class A delay lines. However, this is not the case. The reason is that in Class B delay lines, it is necessary to load an entire packet or segment of a packet into the delay line before the velocity is adiabatically reduced. Thus, a single Class B delay line cannot continuously accept input data. As a result, its effective bandwidth or throughput is limited.

To avoid the throughput limitation on Class B delay lines, a number p of Class B delay lines can be connected in parallel as shown in Fig. 2(b). The input block, is a $1 \times p$ optical switch and the output block is a $p \times 1$ optical switch. This structure approximates a RAM because packets can be output from the buffer in any order. The chip area of the delay lines in a Class B buffer, using the structure in Fig. 2(b), is approximately the same as for a Class A buffer [Fig. 2(a)] with the same capacity. Thus, it would appear that Class B buffers offer no advantages over Class A buffers. However, the capacity of each individual delay line in Fig. 2(b) is a factor of p smaller than the delay line in Fig. 2(a). We will show later in this section that the power dissipation in line buffers is proportional to the square of the capacity. Therefore, the Class B buffer dissipates a factor of p less power than the Class A buffer.

In an ideal slow-light delay line, the minimum length of a stored bit of data is approximately one wavelength λ [18]. This is much smaller than could be achieved with fiber delay lines. To minimize the chip area of slow-light waveguides, it is desirable to arrange the waveguide so that maximum use is made of the available chip area. This could be achieved by folding the waveguide, as shown in Fig. 3. If we assume spacing between waveguides of 5λ, the chip area occupied by one stored bit of data is $5\lambda^2$. In Section III, we compare practical devices with this ideal figure and also with projected figures for CMOS embedded dynamic RAM (eDRAM).

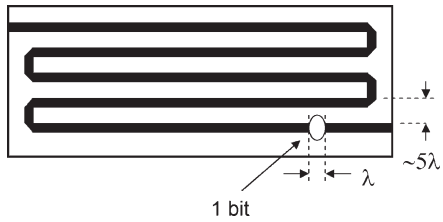

Fig. 3. Planar structure for slow-light waveguide delay line.

For ON-OFF keyed (OOK) optical data and direct detection, the number of bits that can be stored in a delay line is limited by the delay-bandwidth product [18]. The capacity $N_{\rm bit}$ of a slow-light waveguide, measured in bits of stored data, is [18]

$$N_{\rm bit} = T_{\rm S} B_{\rm e} = \frac{L}{L_{\rm bit}} = \frac{L B_{\rm e}}{\nu_{\rm g}} \tag{1}$$

where $T_{\rm S}$ is the delay, or storage time, $B_{\rm e}$ is the baseband (electrical) bandwidth (i.e., the bit rate), and L is the length of the delay line, $\nu_{\rm g} = c/n_{\rm g}$ is the group velocity, $n_{\rm g}$ is the group index, c is the speed of light in air, and $L_{\rm bit}$ is the length of a single bit in the delay line.

C. Attenuation Limitations

We show in this paper that attenuation is a key limiting factor that affects the performance of optical delay lines. For EIT waveguide delay lines, there are two components of attenuation: One is caused by the EIT medium [17], and the other is caused by residual (primarily scattering) losses in the waveguide. Following the analysis in [17] of the attenuation in the EIT medium, it can be shown that the power attenuation per unit length α in an EIT medium at a detuning frequency of $\delta\omega = \omega - \omega_0$ is

$$\alpha(\delta\omega) \cong \frac{1}{\nu_{\rm g}\tau}\left(1 + \frac{(\delta\omega)^2}{\Omega^2}\right) + \alpha_{\rm R} \tag{2}$$

where $\omega = 2\pi\nu$ is the optical radian frequency, ω_0 is the center frequency of the EIT-induced optical passband, $\tau = t_\alpha/\ln(2)$ is the absorption time (i.e., the time taken for a delayed pulse to be attenuated by a factor of e^{-1}), t_α is the -3-dB absorption time defined in [17], $\alpha_{\rm R}$ is the residual waveguide attenuation, and the group index is $n_{\rm g} \cong n\Omega_p^2/\Omega^2$, where Ω_p is the material-dependent plasma frequency and Ω is the half-width of the EIT transmission band [17]. In practical EIT delay lines, where $\nu_{\rm g} \ll c$, the attenuation is dominated by the first term in (2). Therefore, in the present analysis, we put $\alpha_{\rm R} \cong 0$.

Like EIT devices, the power attenuation per unit length in CWR delay lines is inversely proportional to the group velocity [39]. There is some debate over the exact form of the attenuation versus group-velocity relationship for PC waveguides operated near the photonic band edge, with reported dependencies varying from an inverse square [40] to an inverse square root [41] relationship. Unlike EIT devices, the attenuation, in

the passband, of CWR delay lines is not strongly dependent on the optical frequency [42]. In the present paper, we model the attenuation α in CWR delay lines, PC waveguides, and fiber delay lines as follows:

$$\alpha = \frac{c\alpha_{\rm I}}{n\nu_{\rm g}} + \alpha_{\rm R} = \frac{1}{\tau\nu_{\rm g}} + \alpha_{\rm R} \tag{3}$$

where n is the average refractive index, $\alpha_{\rm I}$ is the intrinsic attenuation of the slow-light medium, $\alpha_{\rm R}$ is the residual waveguide loss and coupling loss, and τ is the absorption time, given by $\tau = n/c\alpha_{\rm I}$. For $\nu_{\rm g} \ll c$ and for practical device lengths, it is reasonable to assume that $\alpha_{\rm R} \cong 0$.

To optimize the performance of slow-light optical delay lines, it is necessary to ensure that the width of the passband of the delay line is matched to the signal bandwidth [18]. We explore this issue in more detail in the next section. But for the present, we assume that the bandwidth of the delay line is equal to or larger than the signal bandwidth. With this assumption, we can find the attenuation at band center $\delta\omega = 0$ over the length of a single stored bit in a delay line from (1)–(3):

$$\alpha L_{\rm bit} = \frac{1}{\tau B_{\rm e}}. \tag{4}$$

Equations (1)–(4) enable comparisons between the capabilities of different delay-line buffers. For example, consider EIT, PC, CRW, and fiber-delay-line buffers. The absorption time in a QD EIT medium is on the order of $\tau = 0.6$ ns [17]. Waveguide-based delay lines (including PC slow-light waveguides, CRW slow-light waveguides, and integrated nonslow-light silica waveguides [34]) with intrinsic attenuations of 0.05 and 0.5 dB/cm have absorption times of $\tau \approx 6$ and 0.6 ns, respectively. We now compare these with the absorption time of a fiber-delay-line buffer using the staggered delay-line architecture, as shown in Fig. 1(c). If each fiber delay line has a delay of 250 ns (equal to one IP packet at 40 Gb/s), and if the attenuation of each crosspoint in Fig. 1(c) is 0.5 dB, then the loss through the buffer is 0.5 dB/packet, which corresponds to an effective absorption time of $\tau = 8$ μs. This absorption time is around four orders of magnitude larger (i.e., the loss is four orders of magnitude smaller) than in EIT and CRW, and PC buffers.

To further illustrate the differences in attenuation between EIT, CRW, PC, and fiber-delay-line buffers, we consider buffers that provide 10 μs of delay (e.g., $N_{\rm bit} = 400$ kb or 40 IPv4 packets at a bit rate of 40 Gb/s). A fiber-delay-line buffer of this capacity, constructed using the configuration in Fig. 1(c), would require 2 km of fiber and 40 crosspoints. With an attenuation of 0.5 dB/crosspoint, the total attenuation in the buffer would be 20 dB or 5×10^{-5} dB/bit. If the buffer is realized as a CRW or PC device, with a residual attenuation of 0.05 dB/cm, the total attenuation at the center of the passband would be 7200 dB or 0.018 dB/bit. If the buffer line is an EIT device using semiconductor device using QDs with $\tau = 0.6$ ns, then the attenuation at the center of the passband would be 72 000 dB, or 0.18 dB/bit. Clearly, fiber-based buffers offer much lower loss than any of the planar waveguide alternatives.

D. Power Dissipation

In general, it is possible to circumvent losses in delay lines by including optical gain in the system. However, optical gain introduces noise and requires additional power dissipation [25], [43]. In the Appendix, we derive expressions for the total dissipated power, the dissipated energy per bit of stored data in a lossy delay-line buffer, and the maximum achievable capacity of a delay line. These results assume that any pump power required to create EIT in the slow-light medium is negligibly small. The expressions for dissipated energy per bit represent lower bounds on the energy dissipation associated with storing a bit of data in an optical delay line. For a given attenuation in the slow-light waveguide, it is not possible to reduce the energy below these bounds.

In the Appendix, it is shown that in a delay line that uses a distributed optical gain to overcome waveguide losses, the dissipated energy $E_{\rm bit}$ per bit is

$$E_{\rm bit} = \frac{n_{\rm sp}h\nu\alpha L}{\eta}\left[4\alpha L\,\mathrm{SNR} + \frac{B_{\rm gain}}{B_{\rm e}}\right] \tag{5}$$

and the signal output power $P_{\rm out}$ of the delay line is

$$P_{\rm out} = 4(\alpha L)^2 B_{\rm e} n_{\rm sp} h\nu \mathrm{SNR} \tag{6}$$

where $n_{\rm sp}$ is the spontaneous-emission factor, h is Plank's constant, ν is the optical frequency, η is the quantum efficiency of the gain medium, SNR is electrical signal-to-noise ratio after detection, and $B_{\rm gain}$ is the emission spectral width of the gain medium.

The maximum capacity of a delay line reaches its maximum value $N_{\rm bit-max}$ when the output power is equal to the saturation power $P_{\rm sat}$ of the gain medium in the delay line. Thus, if we put $P_{\rm out} = P_{\rm sat}$ and $N_{\rm bit} = N_{\rm bit-max}$ in (1), (3), (5), and (6), we obtain

$$E_{\rm bit-max} = \frac{P_{\rm sat}}{\eta B_{\rm e}}\left[1 + \frac{\tau B_{\rm gain}}{4N_{\rm bit-max}\mathrm{SNR}}\right] \tag{7}$$

and

$$N_{\rm bit-max} = \frac{\tau}{2}\sqrt{\frac{P_{\rm sat}B_{\rm e}}{n_{\rm sp}h\nu\mathrm{SNR}}} \tag{8}$$

where $E_{\rm bit-max}$ is the energy per bit in a delay line at maximum capacity. Note that if $N_{\rm bit-max} \gg \tau B_{\rm gain}/4\mathrm{SNR}$, (7) reduces to

$$E_{\rm bit-max} \cong \frac{P_{\rm sat}}{\eta B_{\rm e}} \cong \frac{4n_{\rm sp}h\nu(\alpha N_{\rm bit}L_{\rm bit})^2\mathrm{SNR}}{\eta}. \tag{9}$$

Equation (8) shows that the saturation-power-limited maximum capacity of all delay lines is proportional to the absorption time. An important implication of this is that the maximum capacity of optical-fiber-delay-line buffers is much larger than the maximum capacity of slow-light and other planar-waveguide buffers. For example, a fiber-delay-line buffer with $\tau = 8$ μs has a maximum capacity about four orders of magnitude larger than slow-light buffers with τ values around 1 ns. Equation (9) shows that the energy per bit increases with the square of the capacity $N_{\rm bit}$.

E. Bandwidth Limitations

In practical slow-light delay lines, pulse spreading due to dispersion can have a significant impact on the storage capabilities of the delay line [17], [44]. In addition, frequency-dependent signal attenuation affects the capacity [17], [18] and power dissipation [25]. The useful bandwidth of an EIT delay line is limited by two effects. The first of these is a compression of the -3-dB optical bandwidth of an EIT delay line as the length is increased, and the second is caused by dispersion in the EIT medium. We consider both of these effects in the following paragraphs.

Like any chain of cascaded passband filters, the bandwidth of an EIT slow-light device is compressed as the transmission length of the device increases [18]. The -3-dB optical bandwidth $\omega_{-3\ \mathrm{dB}}$ of an EIT delay line as a function of its length L is obtained by putting $\alpha(\omega_{-3\ \mathrm{dB}}/2) - \alpha(\omega_0) = \ln(2)/L$. Thus, from (2), we obtain

$$\left(\frac{\omega_{-3\ \mathrm{dB}}}{\Omega}\right)^2 = \frac{4\ln(2)\nu_{\mathrm{g}}\tau}{L}. \tag{10}$$

Equation (10) shows that the bandwidth $\omega_{-3\ \mathrm{dB}}$ decreases as the length L of the delay line increases (i.e., as the number of stored bits increases). For maximum capacity, the optical bandwidth of the delay line is set to be equal to the optical bandwidth $2B_{\mathrm{e}}$ of the modulated optical signal [18]. Thus, we put $\omega_{-3\ \mathrm{dB}} = 4\pi B_{\mathrm{e}}$ in (10) and substitute (1) and (2) to obtain the amplitude-bandwidth-limited length $L_{\mathrm{amp}}^{\mathrm{EIT}}$ as follows:

$$L_{\mathrm{amp}}^{\mathrm{EIT}} = \frac{4\pi^2 c}{\ln(2) n\tau\Omega_p^2} N_{\mathrm{bit}}^2. \tag{11}$$

Khurgin [17] has shown that the dispersion limit on optical bandwidth translates into the following dispersion-limited length $L_{\mathrm{dis}}^{\mathrm{EIT}}$:

$$L_{\mathrm{dis}}^{\mathrm{EIT}} = \frac{6cB_{\mathrm{e}}}{n\gamma^3\Omega_p^2} N_{\mathrm{bit}}^2 \tag{12}$$

where $\gamma \sim 0.3$ is a function of the modulation format. Note that the amplitude- and dispersion-limited lengths are equal when $B_{\mathrm{e}} = 2\pi^2\gamma^3/3\ln(2)\tau$.

The useable bandwidth of CRW delay lines is limited by a third-order dispersion. Following the analysis in [17], it can be shown that the dispersion-limited length $L_{\mathrm{dis}}^{\mathrm{CRW}}$ of a CRW delay line is approximately

$$L_{\mathrm{dis}}^{\mathrm{CRW}} \approx \frac{c\pi^{3/2}}{\Omega_{\mathrm{CRW}}} N_{\mathrm{bit}}^{3/2} \tag{13}$$

where Ω_{CRW} is a measure of the "strength" of the CRW resonance [17].

We now show some examples of calculated energy per bit and maximum capacity using the above relationships. Fig. 4 shows the maximum capacity $N_{\mathrm{bit-max}}$ and the energy per bit E_{bit} as a function of the saturation power P_{sat} in the gain medium. The capacity and energy per bit are given for EIT and CRW delay lines with absorption times of $\tau = 0.6$ ns (i.e., CRW delay line with an intrinsic attenuation of 0.5 dB/cm), CRW delay lines and low-loss integrated silica waveguides [34]

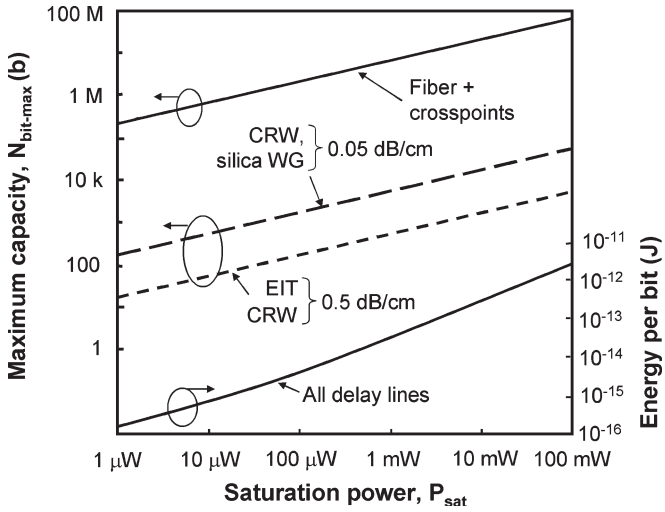

Fig. 4. Capacity and energy per bit at 40 Gb/s versus saturation power.

with an absorption time of $\tau = 0.6$ ns (intrinsic attenuation of 0.05 dB/cm), and a buffer comprising fiber delay lines and crosspoints with 0.5 dB of loss per crosspoint. The energy per bit in the fiber-based buffers does not include the drive power to the crosspoints. In Fig. 4, we have used a wavelength of 1.55 μm, an excess bandwidth B_{excess} of 5 THz, and an electrical bandwidth B_{e} (bit rate) of 40 GHz. For all curves in Fig. 4, we assume a SNR of 20 dB, an ideal gain medium with $n_{\mathrm{sp}} = 1$, and a quantum efficiency η of 100%. These latter assumptions correspond to an ideal device. Therefore, the energies in Fig. 4 represent lower limits on the energy per bit. If the SNR is 30 dB or more (a requirement that would be realistic in a system with cascaded multiple optical delay lines), the energies per bit would be at least an order of magnitude larger than shown in Fig. 4. The energies will also be larger if the spontaneous-emission factor of the gain medium is greater than unity or if the quantum efficiency is less than 100%. The energy per bit in Fig. 4 increases with the delay-line capacity (i.e., with the saturation power), and is on the order of 1-pJ/b delay lines with a saturation power of 50 mW.

It can be seen from Fig. 4 that if the saturation power in the gain medium is 50 mW, then the maximum capacity of a CRW buffer with an intrinsic attenuation of 0.5 dB/cm is 3.7 kb, which is less than half of an IPv4 packet. Similarly, the maximum capacity of an EIT buffer is less than one-half of an IPv4 packet, which is at least an order of magnitude less than what is required for OPS. We conclude that CRW and EIT slow-light delay lines are not suitable for contention resolution in OPS. For the same saturation power (50 mW), a fiber-based buffer could achieve a capacity of up to 50 Mb, or about 5000 IP packets. However, as pointed out earlier, the fiber-based buffers are bulky.

A key question to ask at this point is: what can be done to increase the capacity of CRW and EIT slow-light delay lines? In principle, this could be achieved by combining a reduction of the loss in the waveguide with compensation of the dispersion, either using a linear dispersion compensation using techniques such as detuned resonators in CRW delay lines [44], or soliton propagation [45]. Unfortunately, since $N_{\mathrm{bit-max}}$

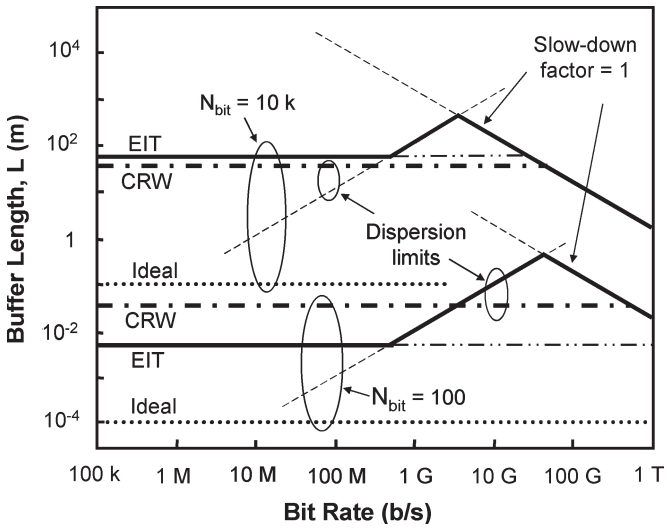

Fig. 5. Buffer length versus bit rate.

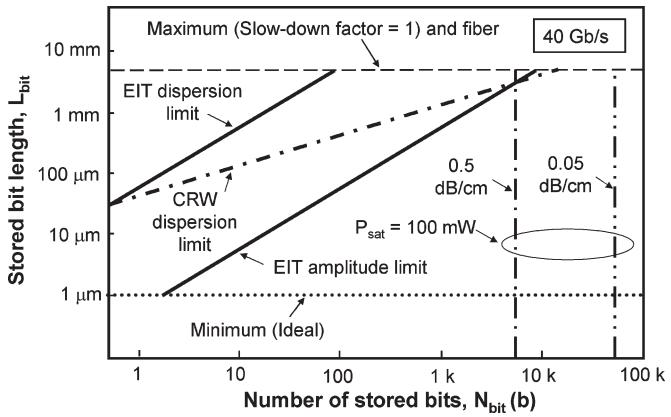

Fig. 6. Bit length versus number of stored bits.

scales with τ, a very large reduction in attenuation is needed to increase the capacity to useful levels. For example, it would be necessary to reduce the intrinsic loss of a CRW waveguide or a nonslow-light silica waveguide to less than 0.001 dB/cm in order to achieve a buffer capacity of around 50 IP packets. If the dispersion in the CRW or EIT devices could be reduced significantly using some form of dispersion compensation, it would be possible to reduce the size of the bits and the length of the delay line. But unfortunately, this would not reduce the power consumption because as the group velocity is reduced, α increases.

Fig. 5 shows the total length of a delay line as a function of the data bit rate for optically amplified CRS and EIT delay lines. The curves are plotted for two buffer capacities: 100 b and 10 kb. Other parameters are the same as for Fig. 4, and the QD material has $\Omega_p = 5 \times 10^{12}$ s^{-1}, and for the CRW, $\Omega_{\mathrm{CRW}} = 4 \times 10^{-13}$ s^{-1} [17]. In Fig. 5, diagonal lines sloping upward to the right represent the bandwidth-dependent dispersion limit in EIT delay lines [see (12)], and the diagonal lines sloping down to the right represent the limit of slow light. On these lines, pulse spreading in the delay line is so large that the slow-down factor becomes equal to unity. The horizontal dot–dash lines in Fig. 5 represent the dispersion limit of CRW devices [see (13)], and the horizontal solid and dashed lines represent the amplitude bandwidth limit of EIT delay lines [see (11)]. The horizontal dotted lines labeled "ideal" represent the fundamental lower limit on buffer length, where the bit length is 1 μm. No slow-light waveguide can be shorter than this.

It can be seen from Fig. 5 that for bit rates below about 500 Mb/s, the dominant limitation on the length L in EIT delay lines is the bandwidth limit. Above 500 Mb/s, the condition $B_{\mathrm{e}} \geq 2\pi^2\gamma^3/3\ln(2)\tau$ is satisfied and the EIT dispersion limit dominates. In principle, above 500 Mb/s, the delay-line length could be reduced by reducing the dispersion through some form of dispersion compensation or by soliton propagation. However, soliton propagation may not be a practical approach because of the very large signal power levels required (on the order of 1 kW in [45]). A reduction in length through dispersion compensation would ultimately be limited by the bandwidth limit (heavy broken lines). At around 3 Gb/s for the 10-kb buffer, and 40 Gb/s for the 100-b buffer, pulse spreading due to dispersion increases to a point where the slow-down factor approaches unity. This indicates that dispersion compensation will probably be required if EIT devices in QDs are to become viable at 40 Gb/s and above. Note that the energy per bit as given in (7) is independent of the length of the delay line. Therefore, the power consumption of a buffer is not improved by introducing dispersion compensation. However, dispersion compensation will, in general, permit the length of the delay line to be reduced.

Fig. 6 shows the length L_{bit} of a stored bit as a function of the number of stored bits for EIT, CRW, and fiber delay lines at a bit rate of 40 Gb/s. The upper limit on L_{bit} is around 5 mm, and it occurs when the slow-down factor is unity (i.e., roughly equal to the bit length in an optical fiber). The lower limit on L_{bit} is around 1 μm (i.e., approximately one wavelength in the waveguide medium). The solid lines in Fig. 6 are the dispersion-limited and amplitude-limited bit lengths for EIT delay lines, and the diagonal dot–dash line is the dispersion-limited bit length for a CRW delay line. Note that the maximum capacity of a CRW delay line without dispersion compensation is around 10 kb, and the maximum capacity of an EIT delay line without dispersion compensation is around 100 b. The dispersion compensation would lead to a reduced waveguide length, but the maximum capacity would still be limited by the saturation power constraint to around 10–100 kb, depending on the waveguide loss. The saturation power constraint is shown in Fig. 4 as vertical dot–dash lines, for a saturation power of 100 mW and for waveguide losses of 0.5 and 0.05 dB/cm.

III. Optical and Electronic Buffers

We now compare the optical-delay-line buffers with electronic buffers in terms of energy dissipation and overall physical size. In this section, we will set aside issues relating to the power dissipation and physical size of O/E and E/O conversions. In Section V, we extend the analysis to include O/E and E/O conversions.

We focus here on a comparison between optical buffers and eDRAM. There are some concerns about whether the latency of DRAM may become an issue in high-speed routers [26].

TABLE I
POWER AND ENERGY DISSIPATION FOR 40-Gb/s DELAY-LINE BUFFERS

				Buffer Delay per Port (Capacity)								
				100 ns (4 kb)			10 μs (400 kb)			1 ms (40 Mb)		
		τ (ns)	SNR (dB)	E_{bit}	Power	Area (cm^2)	E_{bit}	Power	Area (cm^2)	E_{bit}	Power	Area (cm^2)
Slow Light	Class A	6.0	30	0.2 pJ	8 mW	0.5	-	-	-	-	-	-
		0.6	30	-	-	-	-	-	-	-	-	-
	Class B (p =100)	6.0	30	4.3 fJ	0.2 mW	5.0	**20 pJ**	**80 mW**	**50**	-	-	-
		0.6	30	0.2 pJ	8 mW	5.0	-	-	-	-	-	-
Fiber		-	30	2x10^{-3} fJ	80 nW	20 m (length)	1.1 fJ	44 pW	2 km (length)	1 pJ	40 mW	200 km (length)
CMOS				0.1 pJ	8 mW	2x10^{-5}	0.1 pJ	8 mW	2x10^{-3}	0.1 pJ	8 mW	2x10^{-1}

However, with shared bus architectures, there is less of a problem [46]. We note that latency is also an issue in optical buffers [31]. Projections to 2018, based on the 2005 International Technology Roadmap for Semiconductors [23] call for eDRAM with an intrinsic read/write energy of 4×10^{-17} J/bit for each cell and a read/write cycle time of 0.2 ns. The projected cell area is 0.0075 μm^2. For comparison, an ideal slow-light waveguide has a cell area per wavelength of 5 λ^2, or 5 μm^2 at $\lambda = 1$ μm. With an array area efficiency of 60% [23], the projected ideal storage density of eDRAM is approximately 100 Tb/m^2, which is three orders of magnitude larger than the storage density of around 100 Gb/m^2 in an ideal slow-light delay line. The storage density of eDRAM is around five orders of magnitude larger than in an EIT or CRS delay line with a capacity of 100 bits, where the bit length is around 100 μm. In practical eDRAM and slow light buffers, the storage density will be lower than these ideal values, but it is clear that on the basis of chip area, optical delay lines are not competitive with electronic buffers.

The read/write energy of an eDRAM is orders of magnitude larger than the intrinsic read/write energy of a single cell. This is because a significant amount of energy is consumed by a range of on-chip devices and interconnections as well as I/O interfaces. State-of-the-art 90-nm CMOS eDRAM chips operate at power levels in the range of 40 mW [47]. For our comparisons with optical buffers, we assume that the total power consumption of 20-nm feature-size CMOS eDRAM together with 40-Gb/s I/O interfaces (which dominate the power consumption) will be around 8 mW or 0.2 mW/Gb/s. For the purposes of our comparisons, we assume here that the overall storage density (not including the I/O interfaces) is 2 Tb/m^2 (i.e., 2% of the ideal storage density given in the previous paragraph).

Table I compares the energy per bit, power dissipation, and chip area of optically amplified CRW slow-light optical buffers, fiber buffers, and CMOS eDRAM for buffers with capacities of 4 kb (100-ns delay), 400 kb (10-μs delay), and 40 Mb (1-ms delay), all at a bit rate of 40 Gb/s. Except where noted, the parameters used in Table I are the same as used for Fig. 4. In Table I, we have used an output SNRs of 30 to enable multiple optical devices to be cascaded. Each of the entries in Table I applies to a single buffer of the specified capacity. A router with n ports will require n of these buffers. In all slow-light delay lines and the fiber delay lines in Table I, we have assumed single-wavelength operation. As pointed out earlier, some efficiency in device area could be achieved through wavelength division multiplexing. However, the energy bit and total power dissipation cannot be improved by employing wavelength division multiplexing. The blank entries in Table I indicate buffers that cannot be realized in practice because the calculated power dissipation is impractically large and/or because the bandwidth limitations (10)–(13) are not met.

Slow-light buffers in Table I are divided into two groups: Class A [using the cascaded structure shown in Fig. 2(a)] and Class B [using the parallel structure shown in Fig. 2(b), with a fan-out of $p = 100$]. The Class A and Class B groups in Table I are subdivided into two types. The first type is a CRW/EIT device with an absorption time of $\tau = 6$ ns and the second is a CRW device with an absorption time of $\tau = 0.6$ ns. The length of a bit of stored data in the CRS and EIT devices is limited by dispersion. Thus, we use the dispersion-limited length of a stored bit obtained to calculate the area of the chip.

It can be seen from Table I that the only feasible slow-light realization for 400-kb buffers is a Class B buffer using CRW delay lines with $\tau = 0.6$ ns (attenuation of 0.05 dB/cm). Because of their low loss, fiber buffers have lower power dissipation than slow-light buffers. However, fiber buffers consume a relatively large amount of space. For example, the 400-kbit fiber buffer in Table I requires 2 km of fiber while a slow-light buffer with the same capacity requires 1 cm^2 of chip area. The chip area of the CMOS eDRAM (not including the area of the I/O interfaces) is five orders of magnitude smaller than this. Table I shows that the only feasible optical solution for 40-Mb buffers is the fiber-delay-line-based buffer.

One additional difference between electronic and optical buffering needs to be mentioned. Electronic buffers can readily provide full random-access functionality. On the other hand, the functionality of delay-line buffers is much more limited than electronic buffers, and the number of additional components such as crosspoint switches (see Fig. 1) increases rapidly as the functionality is increased [31]. These components will occupy an additional space and consume an additional power. Finally, it is worth noting that while slow-light delay lines will not always provide a competitive solution for packet buffering, they may provide a viable platform for packet synchronization.

Packet synchronizers [48] require a maximum controllable delay of one packet length (~250 ns at 40 Gb/s), which is more likely to be practically realizable than larger buffers for contention resolution.

IV. Cross Connects

We now consider the power dissipation and scaling properties of cross connects, and compare optical and electronic switch fabrics. An average size IPv4 packet contains around 10 kb of data and has a length of around 250 ns at 40 Gb/s or 63 ns at 160 Gb/s. To ensure efficient use of the switching fabric, the switching time should be small compared with the packet length. Thus, the switching time in high-capacity cross connects needs to be on the order of 1 ns or less. Modern high-capacity routers employ a small amount of buffer memory within the switch fabric in order to improve the efficiency of scheduling in the switch fabric [49]. However, the size of this buffer is much smaller than the buffer on the line cards and is not considered in the analysis below. In principle, all-optical (i.e., optically controlled) header processing and switching is possible in optical packet switching [4], [50]. However, while this idea is intuitively appealing at one level, it is very unlikely to find a place in medium- or large-size packet switches, because management functions in routers, such as IP-lookup [19], require intensive computing. Electronics is the only technology that could conceivably perform these complicated tasks. Therefore, our considerations of cross-connect fabrics are restricted to electronically controlled cross connects.

A variety of electronically controlled optical switch technologies have been proposed for use in OPS. An excellent review of switch technologies for OPS is given in [51]. These technologies include Lithium Niobate crosspoints [29], [52], semiconductor-optical-amplifier (SOA) gate arrays [53]–[55], wavelength routing methods based on TWCs, and AWG routers [56]–[58]. Although not always mentioned in the literature, a key parameter in optical switch technologies is the power dissipation or switching energy. Indeed, many researchers, including the present author, have paid scant attention to power-dissipation issues when describing experimental demonstrations of optical packet switching.

From a power-dissipation point of view, potential advantage of optical cross-connect technologies over electronic cross-connect technologies in OPS is that optical cross connects change state on a packet-by-packet basis rather than a bit-by-bit basis. In electronic cross connects, a basic unit of switching energy is expended at least once for each bit. Thus, if an electronic switch and an optical switch both have the same switching energy and they are used to switch packets containing 10^4 bits, then the electronic cross connect would consume 10^4 times more energy than the optical cross connect. Optimism about the potential of optical cross connects is further bolstered by studies comparing the limiting switching energies of optical and electronic switches [59] which indicate that the energy dissipation in optical switches could ultimately approach the same thermal limit as electronic switches. However, as we show in this section, a major component of the power consumption in large cross connects is the power consumed by the interconnects. Unlike ideal (loss less) optical switches, real interconnects dissipate a quantum of energy for every bit, not just for every packet.

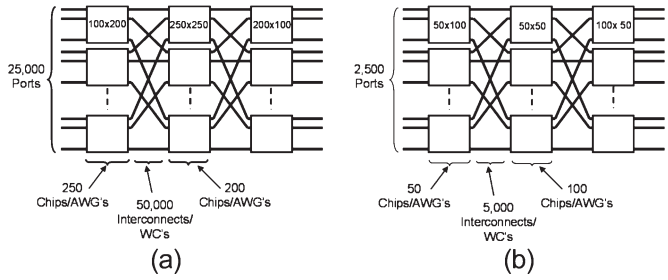

Fig. 7. Clos cross connect architectures at 40 Gb/s (a) 1-Pb/s throughput and (b) 100-Tb/s throughput.

To help clarify the relative advantages of different switch technologies, this section compares the properties of cross connects using AWGs with TWCs, SOA gate arrays, and electronic cross connects. An exhaustive quantitative comparison of all available cross-connect technologies is beyond the scope of this paper. Instead, we have chosen to focus on a few of the most promising options. For the purpose of this comparison, we consider cross connects with nominal capacities of 1 Pb/s and a 100 Tb/s. We consider the cross-connect architectures shown in Fig. 7(a) and (b). Fig. 7(a) represents a 1-Pb/s three-stage Clos 25 000 $\times$ 25 000 cross connect, and Fig. 7(b) is a 100-Tb/s three-stage Clos 2500 $\times$ 2500 cross connect. In both cases, the line rate is 40 Gb/s on each port. The results of our comparison between cross-connect technologies are summarized in Table II, which is explained in detail later in this section.

The most promising cross connects for OPS are AWG-based circuits. For these cross connects, each box in Fig. 7 represents an AWGM and the lines represent single-channel interconnects incorporating rapidly TWCs. A unique advantage of the three-stage Clos architectures in Fig. 7 is that, with only three stages, they can accommodate large port counts. The three-stage Clos architecture is particularly attractive for AWG-based optical cross connects, because the required number of wavelength converters scales linearly with the number of ports, and for a wide range in the number of ports, each input-to-output signal path traverses only four wavelength converters. Thus, while we have not considered every possible cross-connect structure in this paper, we believe that we have used a useful base for comparison.

For SOA-based cross connects, each box in Fig. 7 represents a gate array and the lines between gate arrays represent optical waveguide or fiber connections. The architectures in Fig. 7 can also be used for high-speed electronic switch fabrics. For electronic fabrics, each box in Fig. 8 represents a switch comprising a chip or a small group of interconnected chips, and each line between chips represents a single-channel interconnect.

Chao *et al.* [60] have proposed a switch structure that takes advantage of the intrinsic high bandwidth of the optical cross connects. The incoming packets are split into a number of cells, and these cells are interleaved and optical time multiplexed to a higher bit rate for transmission through an optical switch fabric. Fig. 8 shows how the idea works. The bit rate is increased to 160 Gb/s by a time-division multiplexing, and the optical

TABLE II
(A) TOTAL POWER CONSUMPTION, ENERGY PER BIT, AND CHIP SIZE FOR 1-Pb/s CROSS CONNECT.
(B) TOTAL POWER CONSUMPTION, ENERGY PER BIT, AND CHIP SIZE FOR 100-Tb/s CROSS CONNECT

(a)

		Power			Energy per bit			Area	
External port rate		160 Gb/s	40 Gb/s		160 Gb/s	40 Gb/s		160 Gb/s	40 Gb/s
			Internal port rate			Internal port rate			
			40 Gb/s	160 Gb/s		40 Gb/s	160 Gb/s		
AWG-based cross-connect	SOA WC's	1.43 kW	2.5 kW	2.3 kW	1.43 pJ	2.5 pJ	2.3 pJ	30 cm^2	1.0 m^2
	O/E/O WC's	1.43 kW	2.3 kW	1.53 kW	1.43 pJ	2.3 pJ	1.53 pJ	30 cm^2	1.0 m^2
SOA gate array		10 kW	15 kW	10 kW	10 pJ	15 pJ	10 pJ	0.1 m^2	1.0 m^2
CMOS		2.9 kW	4.6 kW	2.86 kW	2.9 pJ	4.6 pJ	2.86 pJ	30 cm^2	1.0 m^2

(b)

		Power			Energy per bit			Area	
External port rate		160 Gb/s	40 Gb/s		160 Gb/s	40 Gb/s		160 Gb/s	40 Gb/s
			Internal port rate			Internal port rate			
			40 Gb/s	160 Gb/s		40 Gb/s	160 Gb/s		
AWG-based cross-connect	SOA WC's	143 W	250 W	157 W	1.43 nJ	2.5 pJ	2.3 pJ	0.9 cm^2	12.4 cm^2
	O/E/O WC's	143 W	230 W	153 W	1.43 pJ	2.3 pJ	1.53 pJ	0.9 cm^2	12.4 cm^2
SOA gate array		300 W	400 W	300 W	3 pJ	4 pJ	3 pJ	123 cm^2	0.12 m^2
CMOS		290 W	460 W	286 W	2.9 pJ	4.6 pJ	2.7 pJ	0.9 cm^2	12.4 cm^2

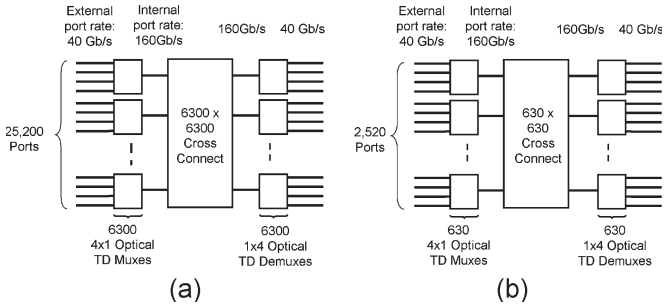

Fig. 8. Cross-connect architectures at 160 Gb/s. (a) 1-Pb/s throughput. (b) 100-Tb/s throughput.

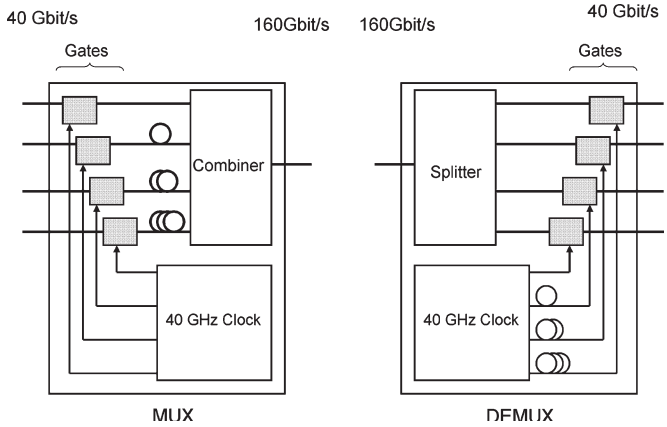

Fig. 9. MUX/DEMUX circuits for Fig. 8.

switch fabrics in the 1-Pb/s cross connect [Fig. 8(a)] and the 100-Tb/s cross connect [Fig. 8(b)] are reduced in size to 6300 × 6300 and 630 × 630, respectively. Incoming packets at 40 Gb/s are split into cells (using circuitry not shown in Fig. 8). The cells at the input ports in Fig. 8 are time-division multiplexed to 160-Gb/s data streams using time-division multiplexers. At the output ports, the data are demultiplexed back to 40 Gb/s.

Fig. 9 shows possible circuits for the multiplexers and demultiplexers for optical switch fabrics. Each multiplexer and demultiplexer uses four high-speed optical gates (either electrically driven or optically driven), a series of three delay lines varying in delay from 0.25 of one bit period to 0.75 of one bit period at 40 Gb/s, and a splitter or combiner. The total length of the delay lines in each multiplexer or multiplexer is 1-bit period. The total length of the delay lines in all multiplexers and demultiplexers in the 1-Pb/s cross connect is 12 500-bit periods at 40 Gb/s or 313 ns. This delay is small compared with the delay required for buffering. To estimate the power consumption of the optical time-division multiplexers and demultiplexers in Fig. 9, we assume that the power consumption is dominated by the optical gates in the multiplexers and demultiplexers. For 1-Pb/s and 100-Tb/s cross connects, there are a total of 50 400 gates and 5040 gates, respectively. Each of which operates at 40 Gsample/s. If we assume a switching energy of 500 fJ/b in each optical gate, the total power dissipation in all gates is 1.0 kW for the 1-Pb/s cross connect and 100 W for the 100-Tb/s cross connect. Electronic multiplexers and

demultiplexers for electronic switch fabrics would use a similar circuitry to Fig. 9. Assuming a 50-fJ/b switching energy in each electronic gate, the total power dissipation in all gates in the multiplexers and demultiplexers is 100 W for the 1-Pb/s electronic cross connect and 10 W for the 100-Tb/s electronic cross connect.

A. Interconnects and Wavelength Converters

Interconnects can become a significant bottleneck in large digital systems [61]. Not only do interconnects dissipate power and occupy significant physical space, both on chip and between chips, they become difficult to manage in large numbers. For the purposes of this paper, we assume that the distances between boards are sufficiently large to require optical interconnects. The interconnects in the electronic cross connects are O/E/O devices and comprise an optical transmitter, a length of fiber or other waveguide, and an optical receiver. In AWG-based cross connects, which require a wavelength converter between each stage, each interconnect is all optical and comprises a wavelength converter and a length of fiber or other waveguide. The TWCs need to be tunable over a large wavelength range, and have a good wavelength registration at each of the channel wavelengths associated with the AWGs.

The outlook for low-power and low-cost E/O/E optical interconnects continues to improve. For example, a high-capacity 250-Gb/s 48-channel interconnect for back plane applications, with a power dissipation of 1.5 W or 6 mW/Gb/s, has been recently demonstrated [62], and a four-channel 10-Gb/s transceiver using 80-nm CMOS and VCSELs has been reported with a power dissipation of 2.5 mW/Gb/s [63]. Near-term projections and power dissipation in optical interconnects [64] suggest that power dissipations below 2 mW/Gb/s are achievable, and it is reasonable to expect that even a lower dissipation will be achievable with future generations of CMOS and other electronic technologies. The design of low-power interconnects involves a range of compromises between design parameters, including receiver sensitivity, receiver power dissipation, and optical power level [64]. For the purposes of this paper, we base our estimates of interconnect performance on the optimized designs presented in [64], with our own projections to 22-nm CMOS technology (6-mW consumption at 6 Gb/s), and scaled up to 40- and 160-Gb/s data rates. If we scale on the basis of constant energy per bit, the power dissipation is 40 mW at 40 Gb/s and 160 mW at 160 Gb/s. On the other hand, if we scale to reduced device feature size (and hight bit rate), assuming constant dissipated power, the energy per bit is much smaller. We steer a middle course between the two extremes and assume intermediate power dissipations of 16 mW (23 mW for tunable) at 40 Gb/s and 50 mW (57 mW for tunable) at 160 Gb/s.

Fig. 10(a) summarizes the power levels and energies per bit for our interconnect model at 40 Gb/s. The model is based on an externally modulated interconnect with a 6-dB link loss (including the modulator). The receiver sensitivity is 40 μW, or 1 fJ/b at 40 Gb/s, the laser output power is 160 μW, the total power consumed by the transmitter is 6 mW (150 fJ/b), and the total power consumed by the receiver is 10 mW (250 fJ/b). In addition to the power components considered in [64], we have included the power consumed by the laser. Note that the power dissipation is dominated by dissipation in the receiver [64].

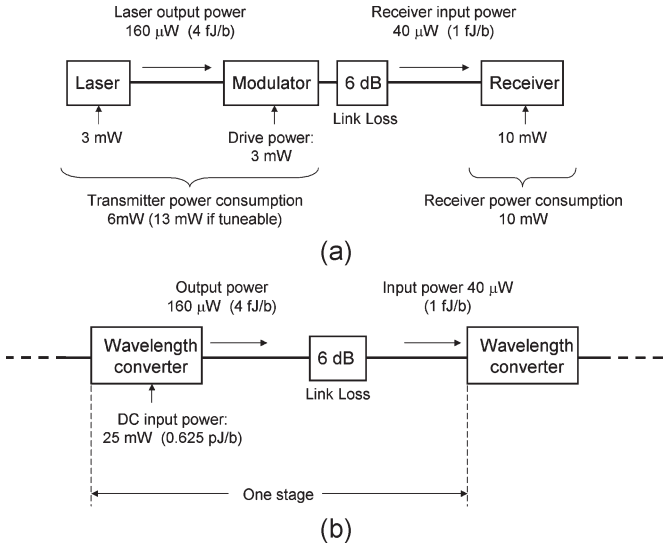

Fig. 10. (a) E/O/E interconnect model and (b) TWC model.

We now consider the power requirements of TWCs for use in AWG-based optical cross connects. Potential technologies for TWCs include cross gain and phase modulation in SOAs [65], four-wave mixing in SOAs [66], [67], and O/E/O wavelength converters. Considerable advances have been made in monolithic integrated tunable all-optical wavelength converters [68], but in optical wavelength converters reported to date, the power requirements are typically an order of magnitude larger than the present-day E/O/E interconnects. Dramatic reductions in power consumption will be needed if all-optical wavelength converters are to become competitive with O/E/O wavelength converters.

An O/E/O wavelength converter is similar to the interconnect in Fig. 10(a), but it has the receiver as its input and the modulator as its output, and the transmitter must be tunable. In general, a tunable laser will consume more power than a fixed-wavelength laser. In our calculations, we assume that the supply power for a tunable laser (including the associated tuning control circuitry) is 10 mW. Fig. 10(b) shows one stage of a cascade of wavelength converters and summarizes the drive power levels and the optical signal levels for an optical wavelength converter based on four-wave mixing or cross modulation in an SOA at 40 Gb/s. The link loss is 6 dB. This loss represents the loss of the AWGs in the cross connect. Based on expected detection sensitivity of SOA wavelength converters [65], [66], we assume that the optical signal levels in Fig. 10(b) are the same as in Fig. 10(a). The dc input power to the wavelength converter supplies both a tunable laser and the SOA. We assume a dc input power of 7 mW to the tunable laser [consistent with Fig. 10(a)], and a dc input power of 18 mW to the SOA, which amounts to a total supply power of 25 mW (625 fJ/b). This is more than an order of magnitude lower than the power requirements of current devices, but we see no fundamental barrier to it becoming achievable in the time fame under consideration in this paper. To estimate the power consumption of an SOA-based wavelength converter operating at 160 Gb/s, we

assume a power consumption of 57 mW at 140 Gb/s (407 pJ/b). This is the same power consumption as we have assumed for tunable O/E/O wavelength converters at 160 Gb/s.

As pointed out earlier, the physical space occupied by the interconnect circuitry can dominate the total area of a cross-connect chip. For the purpose of our calculations here, we assume that the linear density of interconnect ports along the edge of the chips is the same as for all cross-connect technologies. We assume that the I/O ports are spaced by 400 μm–or about twice the diameter of an optical fiber–on the edges of the chips (i.e., the linear density of I/O connections along the chip edge is 25 cm^{-1}). Thus, a 250 $\times$ 250 cross connect with interconnect circuitry occupies a chip area of 25 cm^2.

B. AWG-Based Switch Fabrics

AWGs combined with arrays of TWCs provide a flexible and scalable basis for constructing large switch fabrics. A key advantage of the AWG-based switch fabrics is that the AWG is a passive device. Provided it has a low loss, it does not contribute significantly to power dissipation in the cross connects. The switching speed of the AWG-based cross connects is limited by the tuning speed of the wavelength converters. Power dissipation is also largely determined by the wavelength converters. The number of ports in an AWG-based switching fabric can be scaled up by increasing the number of AWGs. This can be achieved by demultiplexing the wavelengths on the incoming fibers and directing one or more wavelengths from each fiber to separate AWGs at the input to the cross connect [69], [70]. If a space switch is included with each AWGM, the scaled up cross connect remains strictly nonblocking [69]. The maximum number of ports that can be achieved in this way is $N = FW$, where F is the number of fibers and W is the number of wavelengths on each fiber. If number of AWGs is F, the space switch at the output of each AWGM has size $F \times F$, and the number of fully TWCs is $2FW$. The maximum number of fibers that can be accommodated is equal to the number of input wavelengths on each AWGM. For example, a 1600-port switch could be constructed using 80 TWCs, forty 40 $\times$ 40 AWGs, and forty 40 $\times$ 40 space switches. At 40-Gb/s per channel, the total throughput of the switch would be 64 Tb/s. In some cases, it may be possible to achieve a reduction in component count using a wavelength multiplexing through the waveband multiplexing of groups of packets [70] and by omitting the space switches. But this comes at a cost of requiring slotted operation and with a loss of strictly nonblocking behavior.

To estimate the power and energy, we ignore the power required for temperature stabilization of the AWGs. In addition, we assume that any unused WCs on unused paths in the cross connect are powered down when they are not in use. Thus, the total power requirement of the AWG-based cross connects is obtained by simply adding the power consumption of all active WCs (i.e., four WCs in each signal path). The results of this calculation are included in Table II.

The total area occupied by all AWGs in the 1-Pb/s cross connect with the 40-Gb/s port rate is approximately 1 m^2 and 30 cm^2 with a port rate of 160 Gb/s. This highlights the extreme importance of high port rates in large cross connects. While the total chip area of 1 m^2 for the 40-Gb/s switch may seem reasonable, the total number of interconnecting fibers is in excess of 153 000—a truly daunting number of interconnects to manage. Even after a factor of four reductions in the number of interconnects in the 160-Gb/s switch, this remains a major practical challenge to be overcome. The number of interconnects in the 100-Tb/s cross connect reduces to approximately 15 000 and 4000 at 40 and 160 Gb/s, respectively.

C. SOA Gate Arrays

There has been a considerable progress in the development of practical SOA gate arrays [71], [72], and a variety of switch architectures incorporating SOA gates have been proposed and demonstrated [73]–[76]. Potentially useful switch architectures include strictly nonblocking broadcast-and-select structures [73]–[75], [77] and rearrangeably nonblocking structures with a reduced gate count [74]. Strictly nonblocking switching fabrics are required for asynchronous OPS, where the packets passing through the switch are not aligned in time.

The number of SOA gates (i.e., controllable SOA gain blocks) required in an $N \times N$ strictly nonblocking switch is $3N(N-1)$, and the number of cascaded SOAs is $2\log_2 N$ [74]. For a rearrangeably nonblocking Benes switch architecture, the number of SOAs is reduced to $2N\log_2(N-1)$ [74]. In some cases, it is possible to achieve slightly lower SOA counts in multistage architectures in which the final stage is replaced by a bank of wavelength converters [77]. The maximum size of SOA gate arrays is limited by the number of SOA gates that can be cascaded. This, in turn, is limited by the noise figure of the devices, the saturated output power of the SOAs, [74]. In addition, requirements on the extinction ratio of the gates affect the lengths of the SOAs, which in turn affects the achievable noise figure. The saturated output power generally increases with increasing on-level drive current, but higher current levels come at the cost of an increased power consumption. Large switching fabrics can be constructed with SOA powers of the SOAs on the order of 100 mW [74].

To estimate the power consumption and the size of SOA-based cross connects, we have modeled SOA-based strictly nonblocking cross connects based on the structures similar to those shown in Figs. 7 and 8. The loss per stage in the distributed gain structure for SOA gate arrays is determined by the -3-dB splitting loss in each stage [74]. In addition, we assume a 0.5-dB coupling loss at the inputs and outputs of each on-chip SOA and an input/output coupling loss of 3 dB for each of the boxes in Figs. 7 and 8. To estimate a lower limit on the power dissipated by these cross connects, we use (5) (Section II) to determine the energy per bit in a signal passing through a cascade of SOAs. As with the calculations for buffers in Section II, we use an SNR of 30 dB. The loss per stage and per input/output is $e^{-\alpha L} = 0.5$ (3 dB) and the quantum efficiency η is 10%. In our calculations for buffers in Section V, we used a value of 1 for $n_{\rm sp}$. For the SOA-based cross connects, we put for $n_{\rm sp} = 2$ to account for the fact that SOA gates are optimized for extinction ratio rather than noise figure [74]. In both sets of calculations, our estimated values of $n_{\rm sp}$ are somewhat optimistic, but are consistent with our broad

objective to use aggressive but plausible predictions of future technologies. The estimated energies and total dissipated power are included in Table II(a) and (b).

In the above calculation, we have assumed that only those SOAs actually passing a signal have power applied. All remaining SOAs are deliberately powered down. Because the ASE level builds up as more gates are traversed, the saturation power level of the gates needs to be larger in the final stages rather than in the earlier stages. From, (6) the output power level at the final SOA stages is a fairly large 300 and 220 mW for the 1-Pb/s and 100-Tb/s cross connects, respectively.

D. Electronic Cross Connects

At the time of writing, the highest capacity commercial single-chip electronic cross connect provides a throughput of 52 Gbits at a line rate of 4 Gb/s and a total power dissipation of 16 W [78]. This power dissipation corresponds to 30 pJ of energy per bit. Szymanski *et al.* [79] have estimated that with 180-nm CMOS technology, single-chip throughputs on the order of 5 Tb/s will be achievable, with power consumption on the order of 26-W per chip (5.2 pJ/b). The chip area would be 343 mm^2. We extrapolate these results to 22-nm CMOS technology and assume in our calculations that single-chip throughputs of 10 Tb/s at line rates of 40 Gb/s (and eventually 160 Gb/s) will be achievable with power dissipations of 10- and 5-W per chip at 40 and 160 Gb/s, respectively, (1 pJ/b and 500 nJ/b) and with active switching areas of each chip on the order of 500 mm^2. Advanced interconnect integration technologies [80] will be the key in achieving this level of throughput.

To estimate the power dissipation of the electronic cross connect, we assume that the 40-Gb/s 250 $\times$ 250 chip (10-Tb/s capacity at 40 Gb/s) and the 40-Gb/s 100 $\times$ 200 chip (4-Tb/s capacity) chip in Fig. 7 operate at with energy dissipations of 1 pJ/b. As with the AWG-based cross connects, only those interconnects between stages that are carrying data are active. The interconnect power consumption is shown in Fig. 10(a).

E. Comparison of Cross-Connect Technologies

Table II(a) and (b) shows the total power consumption, the energy per bit, and the chip size for a 1-Pb/s cross connect and a 100-Tb/s cross connect, respectively. The tables show data for an AWG-based cross connect, SOA gate arrays, and an electronic cross connect. Data are presented for cross connects operating at a port rate (i.e., the bit rate at each port) of 40 and 160 Gb/s. For the 40-Gb/s port rate, the data is further subdivided into: 1) cross connects that are operating at a 40-Gb/s port rate throughout [see Fig. 7(a) and (b)] and 2) cross connects operating with an internal port rate that is multiplexed up to 160-Gb/s port rate [see Fig. 8(a) and (b)].

If the bit rate per port is increased by a factor of four to 160 Gb/s and the number of ports is reduced by the same factor to 6300, as shown in Fig. 8, the number of ports on each AWG is reduced to 80. The total number of AWGs becomes 240, the number of TWCs becomes 25 600, and the power dissipation in the 6250 $\times$ 6250 cross connect falls to 2.5 kW or 2.5 pJ/bit. If we now add the power dissipation associated with the multiplexers and demultiplexers, the total power dissipation is 12 kW or 12 pJ/bit.

It can be seen from Table II that SOA gate arrays are not competitive with the other two technologies. Its power consumption is high, the chip area is large, even at a line rate of 160 Gb/s, and, as shown earlier in this paper, the saturation powers are large. At a line rate of 40 Gb/s, the electronic cross connect has a power dissipation which is approximately twice that of the AWG-based cross connect. The electronic cross connect offers a size advantage. By increasing the line rate from 40 to 160 Gb/s in the 1-Pb/s cross connect using optical time-division multiplexing of cells [60], the AWG-based solution has a power consumption which is similar to the power consumption of the electronic cross connect. Note that a significant component of the power consumption in this calculation comes from the gates in the multiplexers and demultiplexers.

V. Routers

In this section, we bring together our estimates of buffer performance and cross-connect performance to obtain a picture of the scaling properties of routers, considering both optical and electronic realizations of buffers and cross connects. Many different architectures have been proposed for optical and electronic routers. These include routers with input buffering [29], output buffering [81], and routers in which the buffering and switching functions are combined [4], [57], [81]. In addition, some all-optical routers have been proposed that maximize the utilization of the components through judicious use of WDM. It would be impossible to attempt to compare all possible architectures. Instead, we focus on each of the four possible combinations of buffer and cross-connect technology (optical and electronic) considered in the previous sections of this paper. For each of these four combinations, we find the lowest power realization of a 1-Pb/s router and a 100-Tb/s router operating at line rates of 40 Gb/s. We do this for input-buffered routers and output-buffered routers. In all cases, the buffer capacity per router port is 400 kb (i.e., 10 μs of buffering per port), and we assume Class B slow-light buffering as shown in Table I.

To provide a common basis for comparison of all technologies considered here, we consider routers operating with full-reach optical fiber links at the input and output ports. Therefore, the optical power level at each input port is set at the sensitivity-limited level determined by high-performance optical receivers, and the optical power level at each output is set at the level of a standard optical transmitter. Fig. 11 shows how this can be achieved for routers with input buffering and Fig. 12 shows how it can be achieved using the output buffering. Before calculating the total power dissipation in the eight architectures in Figs. 1 and 12, we need to consider three of extra contributions to the total power that have not been considered so far.

A. Low-Noise Preamplifiers

The fiber inputs to the electronic buffers in Fig. 11(a) and (b) and all cross connects in Fig. 12 are fed through a low-noise optical preamplifier. In Figs. 11 and 12(a) and (c), the low-noise amplifier feeds an O/E converter which is identical to the

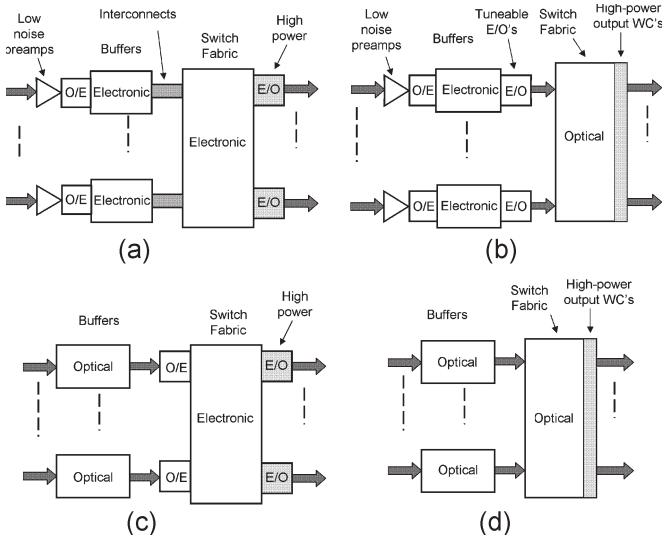

Fig. 11. Input buffered routers.

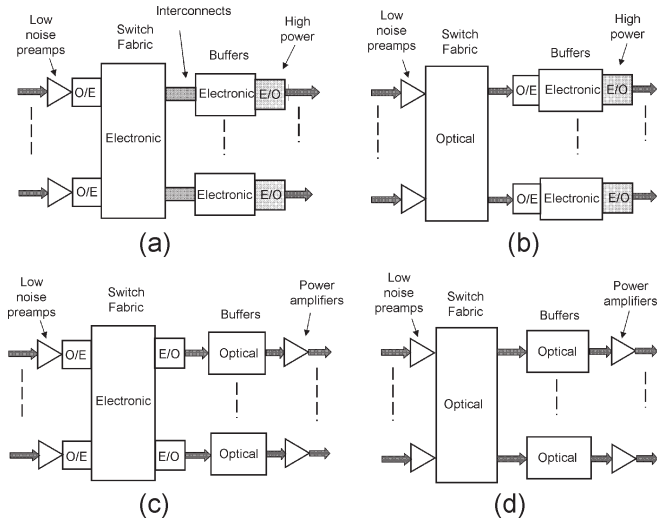

Fig. 12. Output buffered routers.

optical receiver in Fig. 10(a). This O/E converter is optimized for low power dissipation but not for optimum sensitivity [64]. The low-noise optical preamplifiers improve the overall input sensitivity so that it becomes equivalent to the sensitivity of a high-performance optical receiver. In practice, this can be achieved by placing a single optical preamplifier, which will be on each input fiber rather than on each router input port. This single amplifier serves all wavelengths on that fiber. Each amplified fiber is then fed to a wavelength demultiplexer (not shown in Figs. 11 and 12), and each output channel from the demultiplexer is then fed to the input ports shown in Figs. 11 and 12. If we assume that each low-noise preamplifier serves 100 channels and consumes a dc power of 500 mW, then the power consumed per port by these amplifiers is 5 mW.

B. Additional O/E Converters

The interconnects in Figs. 11(a) and 12(a) are part of the cross-connect fabric (see Fig. 7), but we need to consider the extra power consumption of the O/E converters at the inputs of the buffers in Figs. 11(a) and (b), and 12(b). The power consumption of each extra O/E converter is 10 mW [see Fig. 10(a)].

C. Tunable E/O Converters

The tunable E/O converters in Fig. 11(b) replace the input bank of wavelength converters in the optical cross connect. Comparing Fig. 10(a) and (b), it can be seen that this constitutes a power saving of 12 mW per port.

D. High-Power E/O Converters and High-Power WCs

To enable the router to operate in a network environment, the E/O converters at the outputs of the electronic cross connects in Fig. 11(a) and (c) and the electronic buffers in Fig. 12(a) and (b) need to operate as optical line transmitters. These transmitters operate at higher power levels than the transmitters in the interconnects, and are shaded in Figs. 11 and 12. Similarly, the wavelength converters at the outputs of the optical cross connect [shaded in Fig. 11(b) and (d)] need to operate at a higher power level than the internal wavelength converters. We assume here that the power consumption of a line-transmitter-grade wavelength converter is 100 mW. Alternatively, instead of using high-power output stages in Fig. 11, line power amplifiers could be employed on the outgoing fiber links (see Section V-E).

E. Line Power Amplifiers

Line power amplifiers at the outputs of the optical buffers in Fig. 12(c) and (d) boost the output power of the optical buffers to line levels. These amplifiers would generally be placed in the outgoing fibers after the WDM multiplexers (not shown in Fig. 12). We assume that each line amplifier serves 100 channels and consumes a dc power of 10 W. The power consumed per port by this amplifier is 100 mW.

Tables III and IV summarize our analysis of the power dissipation and size of routers, including both buffers and cross connects. Table III is for a 1-Pb/s router with input buffering and Table IV is for a 1-Pb/s router with output buffering. We have restricted the analysis to routers with external and internal port rates of 40 Gb/s. However, we obtain similar conclusions if we include cross connects with 160-Gb/s internal port rates (see Fig. 8). For simplicity, we have not included data for the 100-Tb/s routers, but these data can easily be obtained from the data in Tables I–IV, and the conclusions are similar. Tables III and IV include the four possible combinations of buffer and cross connect, using optical and electronic solutions, as shown in Fig. 11, and list the power and chip area for each combination. The entry labeled "extra power" in the tables corresponds to the additional power consumed by all the additional components highlighted in Figs. 11 and 12. The data in the tables were assembled by taking the power consumption and area data for the buffers and cross connects from Tables I and II, and scaling the data to the number of ports in each case. The optical buffer power dissipation in Tables III and IV includes the power dissipation in wavelength converter required in the $1 \times p$ input switch in Fig. 2(b). The power and area figures for the optical cross connects were based on the AWG-based cross connects using O/E/O-based WCs, because this solution offers

TABLE III
TOTAL POWER CONSUMPTION AND CHIP SIZE FOR 1-Pb/s INPUT-BUFFERED ROUTER

			Buffer							
			Optical				Electronic			
			Buffer	Cross connect	Extra power	Total	Buffer	Cross Connect	Extra power	Total
Cross connect	Optical	Power	2.0 kW	2.3 kW	1.9 kW	6.2 kW	0.2 kW	2.3 kW	2.0 kW	4.5 kW
		Area	125 m^2	1 m^2	-	126 m^2	50 cm^2	1 m^2	-	1 m^2
	Electronic	Power	2.0 kW	4.6 kW	2.0 kW	8.6 kW	0.2 kW	4.6 kW	2.5 kW	7.3 kW
		Area	125 m^2	1 m^2	-	126 m^2	50 cm^2	1 m^2	-	1 m^2

TABLE IV
TOTAL POWER CONSUMPTION AND CHIP SIZE FOR 1-Pb/s OUTPUT-BUFFERED ROUTER

			Buffer							
			Optical				Electronic			
			Buffer	Cross connect	Extra power	Total	Buffer	Cross Connect	Extra power	Total
Cross connect	Optical	Power	2.0 kW	2.3 kW	2.6 kW	6.9 kW	0.2 kW	2.3 kW	2.5 kW	5.0 kW
		Area	125 m^2	1 m^2	-	126 m^2	50 cm^2	1 m^2	-	1 m^2
	Electronic	Power	2.0 kW	4.6 kW	2.2 kW	8.8 kW	0.2 kW	4.6 kW	2.1 kW	6.9 kW
		Area	125 m^2	1 m^2	-	126 m^2	50 cm^2	1 m^2	-	1 m^2

a minimum power. However, we note that similar results are obtained if we use SOA-based WCs.

It can be seen from Tables III and IV that the chip area is dominated by the buffer when optical buffers are employed. In addition, the power dissipation of the slow-light optical buffer is significant. Similarly, the physical size is dominated by the cross connect when electronic buffers are used. While not shown explicitly in Table III, the total buffer power dissipation and size could be reduced if the buffer size on each port could be reduced, but for the reasons given in Section II, this may not be possible. On the other hand, if it is necessary to increase the buffer size significantly beyond 400 kb, the slow-light optical technology clearly cannot offer a viable solution. Fiber-based delay lines would be technically feasible, but very bulky. A 40-Mb slow-light optical buffer would require many orders of magnitude reduction in waveguide losses and/or new dispersion compensation techniques that could permit the 1-μm bit size to be approached. In comparison, the size of electronic buffers in the 1-Tb/s router could easily be increased by orders of magnitude beyond 400 kb, and this would have a minimal impact on the total power dissipation or chip area. Taking these considerations together, electronic buffering appears to offer a significant advantage over optical buffering.

Table III shows that the minimum power dissipation for a 1-Pb/s input buffered router (4.5 kW) is achieved using the AWG-based cross connects (incorporating O/E/O WCs) and electronic buffers. Table IV shows that similar conclusions apply to the output buffered routers. The reason for these small differences in the total calculated power dissipation is the differences in interconnect and wavelength converter power dissipations in the various configurations in Figs. 11 and 12.

The power dissipation in the cross connects is lower in optical cross connects than in electronic cross connects. This is because, AWGs do not consume significant power, but the electronic switching blocks in electronic cross connects do indeed consume power. While the differences between power dissipation in electronic and AWG-based cross connects are small, the minimum power dissipation in complete routers (and minimum chip size) is achieved in routers that combine optical cross connects and electronic buffers.

VI. CONCLUSION

We have investigated the power and chip-size requirements of buffers and cross connects for future high-capacity routers. To do this, we have compared the properties of optical and electronic alternatives. Our analysis is based on aggressive but plausible estimates of optical and electronic device performance, projected out to around 2020. Our broad conclusion is that optical and electronic technologies will both be integral components of future high-capacity routers. However, optical-delay-line buffers do not appear to provide a viable alternative to electronic buffers. Slow-light optical buffers are seriously limited by intrinsic losses and the attendant power dissipation. It is unlikely that these losses can be reduced to a point where slow-light buffers become practical for OPS. Fiber-delay-line buffers are physically large. However, they have much lower intrinsic loss and lower power dissipation than the planar waveguide devices. Fiber delay lines appear to provide the only viable medium for optical delay line buffering in routers.

Electronic buffering offers a number of significant advantages over the optical buffering. We believe that electronic buffers will be the technology of choice in future generations of high-capacity routers. Electronic RAM is attractive because of its small size, low power consumption, its ability to provide full RAM functionality, its scalability to higher capacities, and the

continuing march forward in its capabilities. Given that there does not appear to be a compelling case for moving toward optically transparent OPS technology. We believe that future generations of packet switches will continue to use electronic buffering.

SOA gate switch arrays do not scale easily to the sizes needed for the switching fabric in large routers. AWG-based wavelength-routed optical switch fabrics are attractive because they provide potential low-power switching and good scalability. We estimate that the power consumption in all-electronic cross connects will be marginally larger than in all-optical cross connects and O/E/O cross connects. A key difficulty in building all-optical AWG-based cross connects is the lack of technology for all-optical wavelength conversion. Our calculations in this paper have assumed that viable wavelength converters will emerge, but even if this happens, wavelength converters based on O/E/O conversions might be more competitive.

Based on the data presented here, key technologies that deserve concerted research effort include: 1) very low energy methods for all-optical (transparent) wavelength conversion; 2) low power agile tunable lasers; 3) very low loss slow-light optical waveguides; 4) electronic devices with sub-20-nm features; 5) emerging electronic devices with potentially higher speed and lower energy dissipation than SiCMOS; and 6) semiconductor platforms that support integrated electronics and optoelectronics. Finally, it needs to be pointed out that large routers will never become a reality without small low-power interconnects that can be easily interfaced to optical and electronic chips. Even if very small and very low power interconnects become a reality, the management of the tens of thousands of interconnects is likely to cause significant difficulties.

This paper will clearly not be the last word on this topic. New technologies will emerge and enhancements of existing technologies will inevitably require revisions of the numbers and assumptions used in the calculations. However, when comparing any new technologies for packet switching, a key consideration needs to be power and energy. This paper will have achieved its principle objectives if it encourages researchers to focus their attention on power and energy considerations in optical buffering and switching technologies, and if it helps to stimulate informed discussion and objective debate on the relative merits of electronic and optical routers.

APPENDIX

In this appendix, we derive expressions for the power dissipation in optically amplified delay-line buffers. Fig. 13 schematically shows the main components of input and output power to an optical buffer. The power dissipated $P_{\rm diss}$ in the buffer is

$$P_{\rm diss} = P_{\rm signal,in} + P_{\rm ASE,in} + P_{\rm control} + P_{\rm EIT} + P_{\rm gain/regen} - P_{\rm signal,out} - P_{\rm ASE,out} \quad \text{(A1)}$$

where $P_{\rm signal,in}$, $P_{\rm signal,out}$, $P_{\rm ASE,in}$, and $P_{\rm ASE,out}$ are the signal and ASE input and output powers, $P_{\rm control}$ is the power dissipated by the circuitry that controls the functionality of the buffer, $P_{\rm EIT}$ is pump power required to create EIT or in the slow-light medium, and $P_{\rm gain/regen}$ is the power required to drive any internal amplification or regeneration that is included in the buffer to overcome waveguide losses. The control power $P_{\rm control}$ includes the power consumed by all control logic circuits and by all crosspoint switches, gates, and other optical components that are used to write and read data to and from the buffer. For example, in buffers using recirculation loops [Fig. 2(b)] and feed-forward buffers [Fig. 2(c)] part of $P_{\rm control}$ will drive the crosspoints in Fig. 2(b) and (c). The control power will be a strong function of the particular buffer architecture chosen for a router, and it is therefore difficult to estimate this power across all architectures. Similarly, it is difficult to estimate practical engineering values $P_{\rm EIT}$ at such an early stage in the development of slow-light devices. (In CRW-based delay-line buffers, which do not require EIT [16], $P_{\rm EIT}$ is zero). Here, we assume that $P_{\rm control}$ and $P_{\rm EIT}$ are small compared with $P_{\rm diss}$. In this context, our estimates of the dissipated power $P_{\rm diss}$ in delay-line buffers are lower limits.

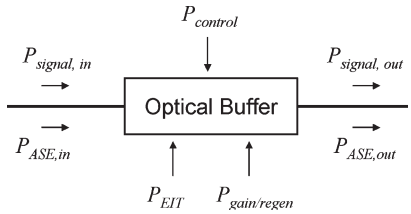

Fig. 13. Power consumption of an optical buffer.

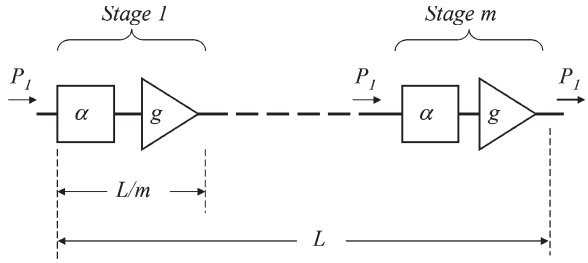

Fig. 14. Amplified delay line.

Our model of an optically amplified lossy waveguide delay line is shown in Fig. 14. The waveguide has a total length of L and comprises m identical stages, each of length L/m. Each stage is modeled by an attenuation block with a loss of $D = {\rm e}^{\alpha L/m}$, where α is the power attenuation per unit length, and a gain block with power gain $G = {\rm e}^{gL/m}$, where g is the gain per unit length. The waveguide has an optical bandwidth of $B_{\rm o}$ and the gain medium has an optical spectral width $B_{\rm gain}$. We include the spectral width of the gain medium into the analysis to account for spontaneous emission emitted by the gain medium outside the bandwidth of the waveguide. This spontaneous emission consumes power and is dissipated by scattering and other effects. We define the excess bandwidth $B_{\rm excess}$ of the gain medium as $B_{\rm excess} = B_m - B_{\rm o} \cong B_m$.

We assume that in each stage, the gain compensates the loss, i.e., $g = \alpha$. The input signal plus noise power into the delay in each signal channel is $P_1 = P_{\rm signal,in} + P_{\rm ASE,in}$, and the output power (signal plus noise) in each channel is $P_{\rm out} = P_{\rm signal,out} + P_{\rm ASE,out}$. Because the internal losses are balanced by the gain, the signal power level at the output of every stage is P_1 and the ASE noise level increases linearly with distance

along the delay line. The total optical signal and ASE power P_n at the input to the nth stage is given by

$$P_n = P_1 + n_{\rm sp}(n-1)(e^{\alpha L/m}-1)h\nu B_{\rm o} \qquad \text{(A2)}$$

where $n_{\rm sp}$ is the spontaneous-emission factor, h is Plank's constant, and ν is the optical frequency. From (A2), it can be seen that the total power increases linearly with $(n-1)$ due to the buildup of ASE noise along the cascade of amplifier stages.

The electrical SNR of the output signal in each channel after detection is given by [82]

$$\text{SNR} = \frac{P_1^2}{4n_{\rm sp}m(e^{\alpha L/m}-1)h\nu B_{\rm e}\left[P_1+n_{\rm sp}m(e^{\alpha L/m}-1)h\nu B_{\rm o}\right]} \qquad \text{(A3)}$$

where $B_{\rm e}$ is the electrical bandwidth of the receiver. We assume here that the electrical bandwidth is equal to the bit rate. Thus, $B_{\rm e} = 1/\tau_{\rm b}$, where $\tau_{\rm b}$ is the bit period.

Our objective here is to explore the lower limits on power dissipation. It is easy to show that the smallest power dissipation occurs when the gain is distributed uniformly along the delay line. For distributed gain, the attenuation $\alpha L/m$ per stage in (A2) and (A3) is small. Under these conditions, (A2) reduces to

$$P_n = P_1 + \frac{n_{\rm sp}(n-1)\alpha L h\nu B_{\rm o}}{m} \qquad \text{(A4)}$$

and the total optical signal and ASE power $P_{\rm out}$ at the output of the delay line is obtained from (A4) with $n = m+1$:

$$P_{\rm out} = P_1 + n_{\rm sp}\alpha L h\nu B_{\rm o}. \qquad \text{(A5)}$$

With (A4) substituted (A3), it becomes

$$\text{SNR} = \frac{P_1^2}{4n_{\rm sp}\alpha L h\nu B_{\rm e}\left[P_1 + n_{\rm sp}\alpha L h\nu B_{\rm o}\right]}. \qquad \text{(A6)}$$

If the optical bandwidth is small, the first term in the square bracket in the denominator of (A6) dominates, and (A6) reduces to the well-known expression for signal-spontaneous beat-noise-dominated SNR

$$\text{SNR}_{\rm sig-sp} = \frac{P_1}{4n_{\rm sp}\alpha L h\nu B_{\rm e}}. \qquad \text{(A7)}$$

We now determine the total optical signal and ASE power dissipated in the delay line. The total signal plus ASE power $P_{\rm d}$ dissipated in the delay line is obtained by summing the power dissipated in all attenuation blocks. Thus

$$P_{\rm d} = \sum_{n=1}^{m} P_n(1-e^{-\alpha L/m}) \approx \sum_{n=1}^{m} P_n \alpha L/m. \qquad \text{(A8)}$$

In addition to the in-band signal and ASE power, an additional optical power of $n_{\rm sp}\alpha L h\nu B_{\rm gain}$ is dissipated by spectral components of spontaneous emission that lie outside the optical bandwidth of the waveguide. Adding this additional dissipated power to $P_{\rm d}$ in (A8), we obtain the total power $P_{\rm dl}$ dissipated in the delay line:

$$P_{\rm dl} = \sum_{n=1}^{m} P_p \alpha L/m + n_{\rm sp}\alpha L h\nu B_m. \qquad \text{(A9)}$$

Substituting (A4) and (A7) in (A9), and for large m, the total optical dissipated power in the amplified delay line becomes

$$P_{\rm dl} = n_{\rm sp}h\nu\alpha L\left[4\alpha L B_{\rm e}\text{SNR}_{\rm sig-sp} + B_{\rm gain}\right]. \qquad \text{(A10)}$$

From (A5) and (A6), the total signal output power from the amplified delay line is

$$P_{\rm out} = 4(\alpha L)^2 B_{\rm e} n_{\rm sp} h\nu \text{SNR}_{\rm sig-sp}. \qquad \text{(A11)}$$

In the present analysis, we assume that the control power $P_{\rm control}$ and the pump power $P_{\rm EIT}$ in (A1) are zero. Thus, the dissipated optical power can be written in terms of the drive power to the amplifier/regenerators and the quantum efficiency η of the gain medium or the regenerators. For equal input and output signal powers and for small input ASE power, the dissipated optical power is $P_{\rm dl} = \eta(P_{\rm gain/regen} - P_{\rm ASE,out})$. Therefore, the total dissipated power is $P_{\rm diss} = P_{\rm dl}/\eta$.

Note that if $B_{\rm gain}$ is small, the second term in the square brackets in (A10) is small compared to the first two terms and the dissipated power is proportional to $(\alpha L)^2$, i.e., proportional to the square of the total loss in the delay line. The energy per bit is the product of the total dissipated power $P_{\rm diss}$ and the bit period $\tau_{\rm b}$. Thus, the energy $E_{\rm bit}$ per bit is given by

$$E_{\rm bit} = \frac{n_{\rm sp}h\nu\alpha L}{\eta}\left[4\alpha L\text{SNR}_{\rm sig-sp} + \frac{B_{\rm gain}}{B_{\rm e}}\right]. \qquad \text{(A12)}$$

ACKNOWLEDGMENT

The author would like to thank P. Farrell, K. Hinton, T. Nirmalathas, G. Pendock, and G. Shen of the University of Melbourne, J. Khurgin of Johns Hopkins University, D. Miller of Stanford University, and G. Epps of Cisco Systems for their stimulating discussions and helpful suggestions.

REFERENCES

[1] R. Parthiban, R. S. Tucker, and C. Leckie, "Waveband grooming and IP aggregation in optical networks," *J. Lightw. Technol.*, vol. 21, no. 11, pp. 2476–2488, Nov. 2003.

[2] A. Vukovic, "Communication network power efficiency—Assessment, limitations, and directions," *Electron. Cool.*, vol. 10, no. 3, pp. 18–24, Aug. 2004.

[3] C. Minkenberg, R. P. Luijten, F. Abel, W. Denzel, and M. Gusat, "Current issues in packet switch design," *ACM SigComm Comput. Commun. Rev.*, vol. 33, no. 1, pp. 119–124, Jan. 2003.

[4] D. Blumenthal, P. Prucnal, and J. Sauer, "Photonic packet switches: Architectures and experimental implementations," *Proc. IEEE*, vol. 82, no. 11, pp. 1650–1667, Nov. 1994.

[5] R. S. Tucker and W. Zhong, "Photonic packet switching: An overview," *IEICE Trans. Commun.*, vol. E82-B, no. 2, pp. 254–264, 1999.

[6] S. Yao, B. Mukherjee, and S. Dixit, "Advances in photonic packet switching: An overview," *IEEE Commun. Mag.*, vol. 38, no. 2, pp. 84–94, Feb. 2000.

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

[8] D. Chiaroni, "Packet switching matrix: A key element for the backbone and the metro," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 7, pp. 1018–1025, Sep. 2003.

[9] C. Qiao and M. Yoo, "Optical burst switching OBS—A new paradigm for an optical internet," *J. High Speed Netw.*, vol. 8, no. 1, pp. 69–84, Mar. 1999.

[10] V. W. S. Chan, "Guest editorial: Optical communications and networking series," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 8, pp. 1441–1443, Aug. 2005.

[11] R. Parthiban, R. S. Tucker, C. Leckie, A. Zalesky, and A. V. Tran, "Does optical burst switching have a role in the core network?" in *Proc. OFC/NFOEC*, 2005, pp. 50–52.

[12] L. V. Hau, S. E. Harris, Z. Dutton, and C. H. Behroozi, "Light speed reduction to 17 meters per second in an ultracold atomic gas," *Nature*, vol. 397, no. 6720, pp. 594–598, Feb. 1999.

[13] P. C. Ku, F. Sedgwich, C. Chang-Hasnain, P. Palinginis, T. Li, T. H. Wang, S. W. Chang, and S. L. Chuang, "Slow light in semiconductor quantum wells," *Opt. Lett.*, vol. 29, no. 19, pp. 2291–2293, Oct. 2004.

[14] C. J. Chang-Hasnain, P. C. Ku, J. Kim, and S. L. Chuang, "Variable optical buffer using slow light in semiconductor nanostructures," *Proc. IEEE*, vol. 91, no. 11, pp. 1884–1897, Nov. 2003.

[15] R. W. Boyd, M. S. Bigelow, N. Lepeshkin, A. Schweinsberg, and P. Zerom, "Fundamentals and applications of slow light in room temperature solids," in *Proc. Annu. Meeting IEEE Lasers and Electro-Opt. Soc.*, 2004, pp. 835–836.

[16] Y. A. Vlasov, M. O'Boyle, H. F. Hamann, and S. McNab, "Active control of slow light on a chip with photonic crystal waveguides," *Nature*, vol. 438, no. 7065, pp. 65–69, 2005.

[17] J. B. Khurgin, "Optical buffers based on slow light in electromagnetically induced transparent media and coupled resonator structures: Comparative analysis," *J. Opt. Soc. Amer. B, Opt. Phys.*, vol. 22, no. 5, pp. 1062–1074, May 2005.

[18] R. S. Tucker, P.-C. Ku, and C. J. Chang-Hasnain, "Slow-light optical buffers: Capabilities and fundamental limitations," *J. Lightw. Technol.*, vol. 23, no. 12, pp. 4046–4066, Dec. 2005.

[19] H. J. Chao, "Next generation routers," *Proc. IEEE*, vol. 90, no. 9, pp. 1518–1558, Sep. 2002.

[20] E. Le Rouzic and S. Gooselin, "160-Gb/s optical networking: A prospective techno-economical analysis," *J. Lightw. Technol.*, vol. 23, no. 10, pp. 3024–3033, Oct. 2005.

[21] R. V. Caenegem, J. M. Martinez, D. Colle, M. Pickavet, P. Demeester, F. Ramos, and J. Marti, "From IP over WDM to all-optical packet switching: Economical view," *J. Lightw. Technol.*, vol. 24, no. 4, pp. 1638–1645, Apr. 2006.

[22] K. Washio, "SiGe HBT and BiCMOS technologies for optical transmission and wireless communication systems," *IEEE Trans. Electron Devices*, vol. 50, no. 3, pp. 656–668, Mar. 2003.

[23] *International Technology Roadmap for Semiconductors*. 2005 Edition. [Online]. Available: http://public.itrs.net/

[24] D. Wischik and N. McKeown, "Part I: Buffer sizes for core routers," *ACM/SIGCOMM Comput. Commun. Rev.*, vol. 35, no. 3, pp. 75–78, Jul. 2005.

[25] R. S. Tucker, "Petabit-per-second routers: Optical vs. electronic implementations," presented at the Optical Fiber Commun. Conf., Anaheim, CA, 2006.

[26] G. Appenzeller, I. Keslassy, and N. McKeown, "Sizing router buffers," in *Proc. SIGCOMM*, Portland, OR, 2004, pp. 281–292.

[27] D. Wischik, "Buffer requirements for high-speed routers," in *Proc. Eur. Conf. Opt. Commun.*, 2005, pp. 23–26.

[28] N. Beheshti, Y. Ganjali, R. Rajaduray, D. Blumenthal, and N. McKeown, "Buffer sizing in all-optical packet switches," presented at the Optical Fiber Commun. Conf., Anaheim, CA, 2006.

[29] J. Spring and R. S. Tucker, "Photonic 2×2 packet switch with input buffers," *Electron. Lett.*, vol. 29, no. 3, pp. 284–285, Feb. 1993.

[30] D. K. Hunter, W. D. Cornwell, T. H. Gilfedder, A. Franzen, and I. Andonovic, "SLOB: A switch with large optical buffers for packet switching," *J. Lightw. Technol.*, vol. 16, no. 10, pp. 1725–1736, Oct. 1998.

[31] S. S. Mughal, R. S. Tucker, and K. Hinton, "A taxonomy of optical buffer architectures," presented at the Conf. Optical Internet, Jeju, Korea, 2006.

[32] H. Yang and S. J. B. Yoo, "All-optical variable buffering strategies and switch fabric architectures for future all-optical data routers," *J. Lightw. Technol.*, vol. 23, no. 10, pp. 3321–3330, Oct. 2005.

[33] E. Burmeister and J. E. Bowers, "Integrated gate matrix switch for optical packet buffering," *IEEE Photon. Technol. Lett.*, vol. 18, no. 1, pp. 103–105, Jan. 2006.

[34] H. Park, E. F. Burmeister, S. Bjorlin, and J. E. Bowers, "40-Gb/s optical buffer design and simulation," in *Proc. Int. Conf. Numer. Simul. Optoelectron. Devices*, 2004, pp. 19–20.

[35] P. Janes, J. Tidstrom, and L. Thylen, "Limits on optical pulse compression and delay bandwidth product in electromagnetically induced transparency media," *J. Lightw. Technol.*, vol. 23, no. 11, pp. 3893–3899, Nov. 2005.

[36] S. Mookherjea and A. Yariv, "Coupled resonator optical waveguides," *IEEE J. Sel. Topics Quantum Electron.*, vol. 8, no. 3, pp. 448–456, May/Jun. 2002.

[37] M. Settle, R. Engelen, T. Karle, M. Salib, A. Michaeli, L. Kuipers, and T. F. Krauss, "Flatband slow light in photonic crystal waveguides," presented at the OSA Topical Meeting Slow and Fast Light, Washington, DC, 2006.

[38] J. Khurgin, "Adiabatically tunable optical delay lines and their performance limitations," *Opt. Lett.*, vol. 30, no. 20, pp. 2778–2780, Oct. 2005.

[39] S. Dubovitsky and W. H. Steier, "Relationship between the slowing and loss in optical delay lines," *IEEE J. Quantum Electron.*, vol. 42, no. 4, pp. 372–377, Apr. 2006.

[40] S. Hughes, L. Ramunno, J. F. Young, and J. E. Sipe, "Extrinsic optical scattering loss in photonic crystal waveguides: Role of fabrication disorder and photon group velocity," *Phys. Rev. Lett.*, vol. 94, no. 3, pp. 033903-1–033903-3, Jan. 2005.

[41] L. O'Faolain, M. Settle, D. O'Brien, and T. F Krauss, "Dependence of loss on group velocity in photonic crystal waveguides," presented at the OSA Topical Meeting Slow and Fast Light, Washington, DC, 2006.

[42] A. Melloni, F. Morichetti, and M. Martnelli, "Linear and nonlinear pulse propagation in coupled resonator slow-wave optical structures," *Opt. Quantum Electron.*, vol. 35, no. 4/5, pp. 365–378, 2003.

[43] J. Khurgin, *Power Dissipation in Slow Light Devices—Comparative Analysis*, 2006, submitted for publication.

[44] ——, "Expanding the bandwidth of slow-light photonic devices based on coupled resonators," *Opt. Lett.*, vol. 30, no. 5, pp. 513–515, Mar. 2005.

[45] J. T. Mok, E. Tsoy, I. C. M. Littler, C. M. de Sterke, and B. J. Eggleton, "Slow gap soliton propagation excited by microchip Q-switched pulses," in *Proc. IEEE Lasers and Electro-Opt. Soc. Annu. Meeting*, 2005, pp. 509–510.

[46] K. Lee, S.-J. Lee, and H.-J. Yoo, "A practical method to use eDRAM in the shared bus packet switch," in *Proc. GLOBECOM*, 2002, pp. 2303–2307.

[47] K. Itoh, "Low-voltage embedded RAMs in the nanometer era," in *Proc. IEEE ICICDT*, 2005, pp. 235–242.

[48] F. Wei and Z. Wen-De, "Noise analysis of photonic packet synchronizer," *J. Lightw. Technol.*, vol. 22, no. 2, pp. 343–350, Feb. 2004.

[49] N. Chrysos and M. Katevenis, "Scheduling in non-blocking buffered three-stage switching fabrics," presented at the Conf. Computer Commun., Barcelona, Spain, 2006.

[50] N. Wada, H. Harai, and F. Kubota, "40 Gb/s interface optical code based photonic packet switch prototype," in *Proc. Opt. Fiber Commun. Conf.*, 2003, pp. 801–802.

[51] J. Gripp, M. Duelk, J. E. Simsarian, A. Bhardwaj, P. Bernasconi, O. Laznicka, and M. Zirngibl, "Optical switch fabrics for ultra-high-capacity IP routers," *J. Lightw. Technol.*, vol. 21, no. 11, pp. 2839–2850, Nov. 2003.

[52] W. L. Ha, R. M. Fortenberry, and R. S. Tucker, "Demonstration of photonic fast packet switching at 700 Mbit/s data rate," *Electron. Lett.*, vol. 27, no. 10, pp. 789–790, May 1991.

[53] R. Fortenberry, A. J. Lowery, W. L. Ha, and R. S. Tucker, "Photonic packet switch using semiconductor optical amplifier gates," *Electron. Lett.*, vol. 27, no. 14, pp. 1305–1307, Jul. 1991.

[54] D. Chiaroni, P. Bonno, O. Rofidal, J. C. Jacquinot, P. Poignant, C. Coeurjoly, F. Fernandez, E. Mestre, J. L. Moncelet, A. Noury, A. Jourdan, T. Zami, A. Dupas, M. Renaud, N. Sahri, D. Keller, S. Silvestre, G. Eilenberger, S. Bunse, W. Lautenschlaeger, and F. Masetti, "First demonstration of an asynchronous optical packet switching matrix prototype for multi-terabit-class routers/switches," in *Proc. 27th Eur. Conf. Opt. Commun.*, 2001, pp. 60–61.

[55] F. Masetti, D. Zriny, D. Verchere, J. Blanton, T. Kim, J. Talley, D. Chiaroni, A. Jourdan, J. C. Jacquinot, C. Coeurjolly, P. Poignant, M. Renaud, G. Eilenberger, S. Bunse, W. Latenschleager, J. Wolde, and U. Bilgak, "Design and implementation of a multi-terabit optical burst/packet router prototype," in *Proc. Opt. Fiber Commun. Conf.*, 2002, pp. FD1-1–FD1-3.

[56] C. Guillemot, M. Renaud, P. Gambini, C. Janz, I. Andonovic, R. Bauknecht, B. Bostica, M. Burzio, F. Callegati, M. Casoni, D. Chiaroni, F. Clerot, S. L. Danielsen, F. Dorgeuille, A. Dupas, A. Franzen, P. B. Hansen, D. K. Hunter, A. Kloch, R. Krahenbuhl, B. Lavigne, A. Le Corre, C. Raffaelli, M. Schilling, J. C. Simon, and L. Zucchelli, "Transparent optical packet switching: The European ACTS KEOPS project approach," *J. Lightw. Technol.*, vol. 16, no. 12, pp. 2117–2134, Dec. 1998.

[57] W. Zhong and R. S. Tucker, "Wavelength routing-based photonic packet buffers and their applications in photonic packet switching systems," *J. Lightw. Technol.*, vol. 16, no. 10, pp. 1737–1745, Oct. 1998.

[58] K. Vlachos, J. Zhang, J. Cheyns, Sulur, N. Chi, E. Van Breusegem, I. T. Monroy, J. G. L. Jennen, P. V. Holm-Nielsen, C. Peucheret, R. O'Dowd, P. Demeester, and A. M. J. Koonen, "An optical IM/FSK coding technique for the implementation of a label-controlled arrayed waveguide packet router," *J. Lightw. Technol.*, vol. 21, no. 11, pp. 2617–2628, Nov. 2003.
[59] T. Palm, L. Thylen, O. Nilsson, and O. Sahlen, "Switch energy requirements for first- and second-order perturbations in asymmetric quantum box optical switches," *IEEE J. Quantum Electron.*, vol. 33, no. 4, pp. 562–565, Apr. 1997.
[60] H. J. Chao, D. Kung-Li, and J. Zhigang, "PetaStar: A petabit photonic packet switch," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 7, pp. 1096–1112, Sep. 2003.
[61] Y. Li, E. Towe, and M. W. Haney, "Special issue on optical interconnections for digital systems," *Proc. IEEE*, vol. 88, no. 6, Jun. 2000, pp. 723–727.
[62] M. E. Ali, G. Panotopoulos, E. de Groot, G. M. Flower, G. H. Rankin, A. J. Schmit, K. D. Djordjev, M. R. T. Tan, A. Tandon, W. Gong, R. P. Telia, B. Law, L.-K. Chia, D. W. Dolfi, and B. E. Lemoff, "Demonstration of a high-density parallel-WDM optical interconnect," in *Proc. IEEE Lasers and Electro-Opt. Soc. Annu. Meeting*, 2004, pp. 459–460.
[63] C. Kromer, G. Sialm, C. Berger, T. Morf, M. L. Schmatz, F. Ellinger, D. Erni, G.-L. Bona, and H. Jackel, "A 100-mW 4 × 10 Gb/s transceiver in 80-nm CMOS for high-density optical interconnects," *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2667–2679, Dec. 2005.
[64] H. Cho, P. Kapur, and K. C. Saraswat, "Power comparison between high-speed electrical and optical interconnects for interchip communication," *J. Lightw. Technol.*, vol. 22, no. 9, pp. 2021–2033, Sep. 2004.
[65] J. T. Ahn, M. L. Lee, H. S. Seo, and K. H. Kim, "Low input power wavelength converter using gain-clamped semiconductor optical amplifier," *Electron. Lett.*, vol. 38, no. 18, pp. 1045–1047, Aug. 2002.
[66] M. A. Summerfield and R. S. Tucker, "Optimization of pump and signal powers for wavelength converters based on FWM in semiconductor optical amplifiers," *IEEE Photon. Technol. Lett.*, vol. 8, no. 10, pp. 1316–1318, Oct. 1996.
[67] T. Simoyama, H. Kuwatsuka, B. E. Little, M. Matsuda, Y. Kotaki, and H. Ishikawa, "High-efficiency wavelength conversion using FWM in an SOA integrated DFB laser," *IEEE Photon. Technol. Lett.*, vol. 12, no. 1, pp. 31–33, Jan. 2000.
[68] M. L. Masanovic, V. Lal, J. A. Summers, J. S. Barton, E. J. Skogen, L. G. Rau, L. A. Coldren, and D. J. Blumenthal, "Widely tunable monolithically integrated all-optical wavelength converters in InP," *J. Lightw. Technol.*, vol. 23, no. 3, pp. 1350–1362, Mar. 2005.
[69] M. C. Chia, D. K. Hunter, I. Andonovic, P. Ball, I. Wright, S. P. Ferguson, K. M. Guild, and M. J. O'Mahony, "Packet loss and delay performance of feedback and feed-forward arrayed-waveguide gratings-based optical packet switches with WDM inputs-outputs," *J. Lightw. Technol.*, vol. 19, no. 9, pp. 1241–1254, Sep. 2001.
[70] S. Bregni, A. Pattavina, and G. Vegetti, "Architectures and performance of AWG-based optical switching nodes for IP networks," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 7, pp. 1113–1121, Sep. 2003.
[71] K. Komatsu, *Hybrid Optical Matrix Gate Switches for Photonic Switching*, 1997.
[72] N. Sahri, D. Prieto, S. Silvestre, D. Keller, F. Pommerau, M. Renaud, O. Rofidal, A. Dupas, F. Dorgeuille, and D. Chiaroni, "A highly integrated 32-SOA gates optoelectronic module suitable for IP multi-terabit optical packet routers," in *Proc. Opt. Fiber Commun. Conf.*, 2001, pp. PD32-1–PD32-3.
[73] J. D. Evankow, Jr. and R. A. Thompson, "Photonic switching modules designed with laser-diode amplifiers," *IEEE J. Sel. Areas Commun.*, vol. 6, no. 7, pp. 1087–1095, Aug. 1988.
[74] R. F. Kalman, L. G. Kazovsky, and J. W. Goodman, "Space division switches based on semiconductor optical amplifiers," *IEEE Photon. Technol. Lett.*, vol. 4, no. 9, pp. 1048–1051, Sep. 1992.
[75] F. Dorgeuiele, L. Noirie, J. P. Faure, A. Ambrosy, S. Rabaron, F. Boubal, M. Schilling, and C. Artigue, "1.28 Tb/s throughput 8 × 8 optical switch based on arrays of gain-clamped semiconductor optical amplifier gates," in *Proc. Opt. Fiber Commun. Conf.*, 2000, pp. 221–223.
[76] T. Chien and W. I. Way, "Dynamic range and switching speed limitations of an N × N optical packet switch based on low-gain semiconductor optical amplifiers," *J. Lightw. Technol.*, vol. 14, no. 4, pp. 525–533, Apr. 1996.
[77] C. Develder, J. Cheyns, M. Pickavet, and P. Demeester, "Multistage architectures for optical packet switching using SOA-based broadcast-and-select switches," in *Proc. Opt. Fiber Commun. Conf.*, 2003, pp. 794–795.
[78] Vitesse, *4.25 Gbps 144 × 144 Crosspoint Switch*, 2005.
[79] T. H. Szymanski, H. Wu, and A. Gourgy, "Power complexity of multiplexer-based optoelectronic crossbar switches," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 13, no. 5, pp. 604–617, May 2005.
[80] A. V. Mule, R. A. Villalaz, P. J. Joseph, A. Naeemi, P. A. Kohl, T. K. Gaylord, and J. D. Meindl, "Polylithic integration of electrical and optical interconnect technologies for gigascale fiber-to-the-chip communication," *IEEE Trans. Adv. Packag.*, vol. 28, no. 3, pp. 421–433, Aug. 2005.
[81] D. K. Hunter, M. C. Chia, and I. Andonovic, "Buffering in optical packet switches," *J. Lightw. Technol.*, vol. 16, no. 12, pp. 2081–2094, Dec. 1998.
[82] N. A. Olsson, "Lightwave systems with optical amplifiers," *J. Lightw. Technol.*, vol. 7, no. 7, pp. 1071–1082, Jul. 1989.

Rodney S. Tucker (S'72–M'75–SM'85–F'90) received the B.E. and Ph.D. degrees from University of Melbourne, Parkville, Victoria, Australia, in 1969 and 1975, respectively.

He is a Laureate Professor with University of Melbourne. He is also also the Research Director of the Australian Research Council Special Research Centre for Ultra-Broadband Information Networks (CUBIN), in the Department of Electrical and Electronic Engineering, University of Melbourne. He has held positions with University of Queensland, University of California, Berkeley, Cornell University, Plessey Research, AT&T Bell Laboratories, Hewlett Packard Laboratories, and Agilent Technologies.

Prof. Tucker is a Fellow of the Australian Academy of Science and a Fellow of the Australian Academy of Technological Sciences and Engineering. In 1997, he was awarded the Australia Prize for his contributions to telecommunications.