Analysis of a Dynamically Wavelength-Routed Optical Burst Switched Network Architecture

Michael Düser, *Student Member, IEEE* and Polina Bayvel, *Senior Member, IEEE*

***Abstract*—The concept of optical burst switching (OBS) aims to allow access to optical bandwidth in dense wavelength division multiplexed (DWDM) networks at fractions of the optical line rate to improve bandwidth utilization efficiency. This paper studies an alternative network architecture combining OBS with dynamic wavelength allocation under fast circuit switching to provide a scalable optical architecture with a guaranteed QoS in the presence of dynamic and bursty traffic loads. In the proposed architecture, all processing and buffering are concentrated at the network edge and bursts are routed over an optical transport core using dynamic wavelength assignment. It is assumed that there are no buffers or wavelength conversion in core nodes and that fast tuneable laser sources are used in the edge routers. This eliminates the forwarding bottleneck of electronic routers in DWDM networks for terabit-per-second throughput and guarantees forwarding with predefined delay at the edge and latency due only to propagation time in the core. The edge burst aggregation mechanisms are evaluated for a range of traffic statistics to identify their impact on the allowable burst lengths, required buffer size and achievable edge delays. Bandwidth utilization and wavelength reuse are introduced as new parameters characterizing the network performance in the case of dynamic wavelength allocation. Based on an analytical model, upper bounds for these parameters are derived to quantify the advantages of wavelength channel reuse, including the influence of the signaling round-trip time required for lightpath reservation. The results allow to quantify the operational gain achievable with fast wavelength switching compared to quasistatic wavelength-routed optical networks and can be applied to the design of future optical network architectures.**

***Index Terms*—Dynamic wavelength allocation, fast circuit switching, optical burst switching (OBS), optical networks, optical packet switching.**

I. Introduction

FUTURE optical packet networks must be able to support not only the increasing traffic volumes, but also the growing diversity of services and dynamically varying traffic patterns. Driven by the increasing traffic in wide area networks (WAN), routers are forced to process throughputs which are likely to increase from hundreds of gigabits per second currently to several terabits per second in the near future. Although quasi-static wavelength-routed optical networks (WRONs) are relatively simple to analyze and design [1], they may not be sufficiently flexible in responding to dynamically varying and bursty traffic loads and service diversity. Conversely, in pure packet networks, functionalities needed for processing of packet header information and forwarding at optical line rate are difficult to scale up, mainly due to the growing mismatch between electronic processor speeds (currently $\sim$ 1–2 GHz) and the optical line rates currently at 10 Gb/s and expected to exceed 40 to 80 Gb/s in the near future. Additionally, the difficulties in achieving all-optical packet networks lie in the complexity of building fast and large, single-stage all-optical packet switches and lack of scaleable optical buffers. Optical burst switching was, therefore, proposed [2], [3] as an attempt to reduce the processing in network nodes needed for packet forwarding. Although there is not a universal definition of optical burst switching, [4] lists a number of characteristics that are inherent to most of the schemes. These are burst size granularity (which lies between packet switching and circuit switching), separation of control information (header) and data (payload), a one-way reservation scheme (for most cases), variable burst length, and no optical buffering.

The architecture presented here uses a two-way reservation mechanism and might be considered to be closer to dynamic circuit switching. Thus, we refer to the conventional OBS to mean one-way reservation schemes. Burst switching is a time-domain multiplexing technique to access fiber or lightpath bandwidth in fractions of the bandwidth of a wavelength channel. Typically, packets are aggregated at the edge of the network in order to reduce the processing overhead and then routed over a bufferless core. Almost all burst-switching schemes proposed to date [2]–[11] assume the use of separate burst header (control) and payload (data) channels, where headers are sent into a bufferless switch network with an appropriately chosen offset time t_{offset} from the data to reserve switch resources for routing the associated data appropriately along the selected path. There is no acknowledgment of path reservation because burst lengths considered are in the range of tens of kilobytes (equivalent to burst durations on microsecond timescales) and, thus, do not allow sufficient time for an acknowledgment of path reservation. Because the core is assumed to be bufferless, bursts may be dropped at any point along the path in case of burst contention—which cannot be resolved, wasting the reserved resources—and this approach, therefore, may not provide the required QoS guarantees.

A two-way reservation scheme has also been suggested [7], but it was assumed that bursts are also sent prior to the receipt of an acknowledgment. Recognizing this as a limitation, schemes [4], [6], [8] have been proposed to provide class-of-service differentiation by offset times, i.e., to assign larger offsets for higher priority traffic. This, however would have the effect of

Manuscript received August 29, 2001; revised December 20, 2001. This work was supported in part by Marconi Communications, Ltd., the HRLD Foundation, and the Royal Society.

The authors are with the Optical Networks Group, Dept. of Electronic and Electrical Engineering, University College London, London WC1E 7JE, U.K. (e-mail: p.bayvel@ee.ucl.ac.uk; mdueser@ee.ucl.ac.uk).

Publisher Item Identifier S 0733-8724(02)02555-0.

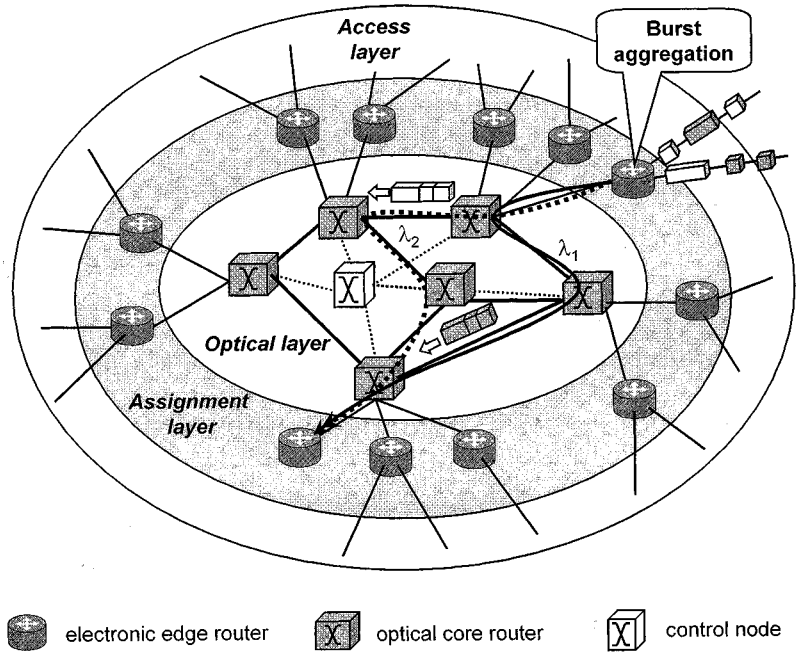


Fig. 1. Architecture of the proposed wavelength-routed optical burst-switched (WROBS) network with burst aggregation at the network edge and wavelength routing in the optically transparent core.

reducing the burst loss for high priority traffic, at the expense of an increase for lower priority bursts, especially for dynamically varying traffic loads. The result is reduced network capacity for acceptable packet loss rates [4]. In OBS networks, lower priority bursts experience loss as a penalty. However, given that each burst may contain a large number of transmission control protocol (TCP)–internet protocol (IP) packets or acknowledgment, each lost burst would affect a number of higher layer connections. Thus, in OBS networks, care should be taken to also minimize the loss of lower priority bursts to prevent this. Even deflection routing would bring little benefit because out-of-order burst arrivals will require large buffers in the receiving edge routers for reordering. Finally, in all of the proposed schemes, wavelengths are assigned on a link-by-link basis, requiring full wavelength conversion at every node, as end-to-end lightpath reservation is difficult because of short offset times and short packets. Hence, wavelengths are not used for routing but simply to increase available transport channel capacity.

In this paper, we propose and analyze an alternative OBS network architecture that requires an end-to-end reservation to satisfy specific service criteria such as latency and packet loss rate (PLR) for bursty input traffic. This architecture, shown in Fig. 1 and termed here wavelength-routed optical burst-switching (WR-OBS), assumes a fast circuit-switched end-to-end lightpath assignment with a guaranteed, deterministic delay, and requires an obligatory end-to-end acknowledgment. The packets are electronically aggregated at the network edge into bursts, according to their destination and class of service (CoS), but with timescale of milliseconds, which is a typical forwarding time of IP routers, making the reservation of resources along the path prior to burst transmission feasible. The aggregation time is strictly determined by the performance parameters such as delay at the edge or the required burst size for the network. At an appropriate point during the aggregation cycle, an end-to-end wavelength channel is requested from a network control node for transmission of the burst between edge routers. Once a free wavelength is found, the aggregated burst is assigned to it and is transmitted into the core network. Its further latency depends only on the propagation delay because buffering operations with associated nondeterministic delays in core nodes are not required. Concentrating all of the processing and buffering within the edge of the network enables a bufferless core network simplifying the design of optical switches or routers/cross connects in the core significantly, which is particularly important for time-critical traffic and cannot be achieved with the currently implemented IP-router infrastructure that provides hop-by-hop forwarding only. This requires, however, that the bit rate at the input to the buffers at edge routers is sufficiently high to form bursts on a millisecond timescale. Following transmission, the wavelength channel is released and can be reused for subsequent connections. The network core can either be considered as a passive core [12] or as a network of fast-reconfigurable optical routers/cross connects, where end-to-end lightpaths or circuits are dynamically set up by the same controller that allocates wavelengths. It is assumed that wavelength conversion in core nodes is not required, because, as previously shown, it brings little benefit to wavelength-routed networks with wavelength agility at the network edge [1]. A centralized network management was assumed in this work as a worst case scenario. A distributed control scheme would be preferred; however, such a scheme relies on synchronization and fast distribution of information on the state of the network.

The aim of this work is to analyze the network performance under which dynamic WR-OBS would bring significant operational advantages and, in particular, in the reuse, utilization

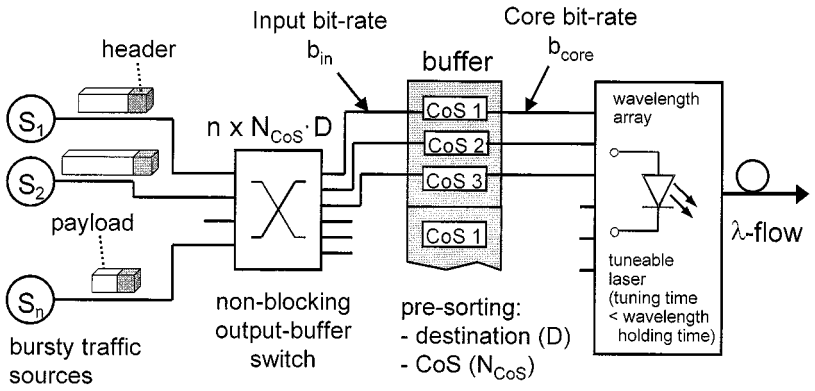


Fig. 2. Edge router model with n bursty traffic inputs, an output-buffered switch with presorting and a wavelength-tuneable laser array for dynamic wavelength allocation. The graph is for $N_{\mathrm{CoS}} = 3$ CoS.

of wavelength channels that are set up only for the required burst transmission time (termed wavelength holding time) and, thus, increased over a much simpler but less adaptable quasi-static logically fully meshed WRON. The calculated values for WR-OBS represent an upper bound for the achievable network parameters, namely the edge delay, bandwidth utilization, wavelength reuse, and idle time, and give design rules on the speed requirements for dynamic routing and wavelength assignment algorithms to make a core network in which resources are assigned dynamically practical. The results can be applied to optimize the design rules of future optical network architectures and quantify the operation regimes that best make use of the static or dynamic network architectures.

The proposed WR-OBS network architecture and the edge router model are described in Section II. The analysis is then separated into two parts. The first studies the burst aggregation process and buffer-induced delays (or edge delays) in edge routers, as a function of different traffic statistics, calculating burst aggregation parameters, including burst size distributions and packet loss rate (PLR). The second part, described in Section III, is dedicated to the study of the optical core and the time bounds for the dynamic wavelength allocation as well as the achievable wavelength reuse, as constrained by the OBS signalling. The analytical model and the results of the core network performance are described in Section IV.

II. Network Architecture and Edge Router Model

A. Network and Edge Router Architecture Assumptions

The proposed edge router setup is shown schematically in Fig. 2, where bursts are aggregated from packets that are electronically presorted according to their destination and CoS and stored in separate queues. After a time-out signal indicates that packets have to be transmitted to meet application specific latency requirements, a wavelength request is sent to a control node, an acknowledgment is received and the buffer content is dynamically assigned to a free wavelength. If a free wavelength channel is not available, packets are not lost and, instead, are stored in edge-router buffers but could incur additional delay. An edge router with dimensions $n \times N_{\mathrm{CoS}} \cdot D$ is considered where n is the number of independent traffic inputs, N_{CoS} represents the number of CoS, and D is the number of destinations. A nonblocking switching architecture is assumed with performance comparable to an output-buffered switch. A combined input and output queueing (CIOQ) switch can achieve such performance when virtual output queueing (VOQ) is used in combination with an internal speed-up [13]. CIOQ switches for arbitrary traffic statistics with a maximum internal speed-up of two were shown to achieve the same performance as an output-queued switch.

The electronic switch, therefore, provides statistical multiplexing because a uniform destination address distribution is assumed. The bit rate b_{in} denotes the aggregated bit rate for the traffic from all n sources directed to a particular destination and requiring the same QoS. Bursts are transmitted from the queue at core bit rate b_{core}, where $b_{\mathrm{core}} > b_{\mathrm{in}}$ and the effect of variations in the ratio $b_{\mathrm{core}}/b_{\mathrm{in}}$ is analyzed in this work.

The edge-router architecture requires several modifications compared to a conventional IP-router architecture. Instead of forwarding packets immediately on the outgoing link, packets are forwarded to buffer queues within the edge router. In considering the buffer size required to implement this operation, it should be taken into account that currently 10-Gb/s line cards are equipped with 128 or 256 MB random access memory (RAM)[1] , sufficient to buffer 107.4 or 214.7 ms of traffic at 10 Gb/s and are, in fact, larger than buffer sizes for burst aggregation considered in Section II-C, where a buffer size of 400 Mb (47.7 MB) is assumed. At 10 Gb/s, this is sufficient to hold 40 ms worth of traffic, a value chosen as a tradeoff between packet loss and delay, although there is no technical limit on the amount of buffering[2] . For nondelay-sensitive traffic, there is an additional advantage of large buffers in this particular architecture; packets are held in the buffer until a free wavelength channel is available, rather than released into the network to be lost on propagation. In the proposed architecture, several edge routers are connected to one optical core router. This simplifies the scalability problem that electronic routers face when their throughput is scaled to the terabit-per-second regime. Assuming a network with $N \approx 100$ edge routers and a core network with $M \approx 20$ core routers, this results in

[1]As of August 2001, Juniper OC-192c linecard: 128 MB, Cisco C-192c linecard: 256 MB.

[2]As of August 2001, Juniper OC-192c linecard: 128 MB, Cisco C-192c linecard: 256 MB.

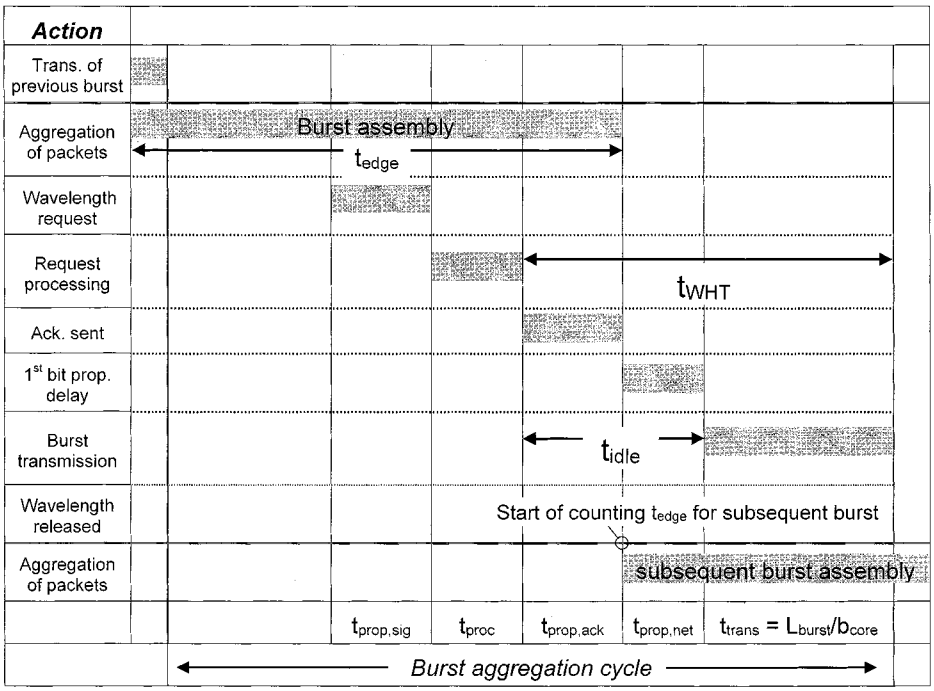


Fig. 3. Timing diagram showing key timing parameters in the burst aggregation and burst transmission processes. The variables used in Fig. 3 to describe the timing of the burst aggregation cycle are as follows. $t_{\rm edge}$ is the maximum delay, i.e., the time the first packet in the buffer spends before the burst is released into the network. $t_{\rm prop,sig}$ is the propagation delay for a control packet sent from the edge router to the network control for wavelength reservation. $t_{\rm proc}$ is the processing time, i.e., time between arrival of control packet and decision on lightpath and wavelength. $t_{\rm prop,ack}$ is the propagation delay between the sending and receiving edge router for sending the acknowledgment for a wavelength reservation $t_{\rm prop,ack} = t_{\rm prop,sig}$, assuming that the control packets take the same route between the sending edge router and the control node. $t_{\rm WHT}$ is the wavelength holding time, i.e., the total time for which a wavelength is reserved. $t_{\rm prop,net}$ is the propagation delay for signal traveling from sending to receiving edge router across the core network. $t_{\rm trans} = L_{\rm burst}/b_{\rm core}$ is the transmission time of the burst. $t_{\rm idle} = t_{\rm prop,ack} + t_{\rm prop,net}$ is the idle time during which the acknowledgment is sent and before the first packet arrives at the receiving edge router.

add–drop traffic on the order of 5 Tb/s per core router for a logically fully meshed architecture with a maximum bit rate per lightpath of 10 Gb/s and higher for the optical core. OBS networks also require new or improved components for dynamic network operation, namely, fast tuneable lasers operating in the C band for DWDM applications [14] and burst-mode receivers with large dynamic range ($>$ 10 dB) and fast clock and data recovery units for bit rates in excess of 10 Gb/s. In core nodes, switching speed and scalability of the number of ports are key design parameters that have to be addressed.

The out-of-band signaling on a separate wavelength requires one control packet per request and acknowledgment, plus several control messages to be sent to core nodes, in the case of switch reconfiguration. Control packets contain information about the origin and destination addresses, the CoS, and the amount of data accumulated when the request was sent to estimate the traffic arrival rate. Assuming control packets of a maximum of 1000 bits to be sent every 10 ms, for a network with $M = 100$ edge routers and three CoS, the capacity of the control network would total approximately 6 Gb/s, or three million requests and acknowledgments per second, compatible with state-of-the-art processors. The requests then have to be processed by the routing and wavelength assignment (RWA) algorithm in the central node. Details addressing both problems can be found in [11], [15].

Different burst aggregation mechanisms have been studied [11]. In this paper, the limited-burst-size scheme (LBS) is used. The control node estimates the traffic arrival rate from the packets accumulated when the request packet is sent and, hence, establishes the burst size by the time the acknowledgment arrives back at the sending edge router. With this approach, it is possible to avoid the release phase required for burst of unknown size. This reduces the overhead experienced, for example, in ATM networks that require a setup and release phase for each connection.

The payload is a concatenation of incoming packets at the ingress and demultiplexed in individual packets at the receiving edge router. The payload would be preceded by a preamble for burst recognition and clock recovery; a few hundred bits have been proven sufficient for this purpose [16]. Although the proposed scheme is protocol independent, it could be integrated in a generalized multiprotocol label switching (GMPLS) environment that supports the establishment of short-lived and on-demand circuits as required for the WR-OBS architecture.

B. Burst Aggregation and Timing Diagrams

The burst aggregation process and most important timings are described in Fig. 3. The burst aggregation cycle can be described as follows. The edge delay $t_{\rm edge}$ is the elapsed time between the time of the arrival of the first bit of the first packet to the buffer queue until the entire burst is released into the network, so that the average queueing delay for all aggregated packets is $t_{\rm edge}/2$. This holds true, however, only in the case of Poisson arrival processes. The arriving packets are aggregated in the buffer until triggered either by a threshold indicating potential buffer

overflow or a timeout signal for delay-sensitive data. This occurs when the wavelength request signalling packet is sent to the control node. The propagation delay for this control packet is $t_{\mathrm{prop,sig}}$.

It is assumed that the signaling packet contains information on the source and destination edge routers, the CoS and the quantity of data in the buffer, required to estimate the wavelength holding time t_{WHT}, defined as the time necessary to empty the buffer and transmit the data between edge routers.

Processing the wavelength request requires time t_{proc}, followed by an acknowledgment packet to be returned to the requesting edge router, with an additional delay $t_{\mathrm{prop,ack}}$. Concurrently with the transmission of $t_{\mathrm{prop,ack}}$, a wavelength channel is reserved, setting the start of t_{WHT}. In parallel, the burst aggregation continues until an acknowledgment from the control node of a confirmed wavelength reservation is received. In this paper, we assume that the burst assembly terminates at the point the acknowledgment packet from the controller reaches the edge router, although alternative schemes have also been analyzed [11]. This allows the burst aggregation to continue in parallel with the processing of the wavelength request, thus decreasing the overall delay although the final burst size would have to be estimated by monitoring the buffer filling statistics. Packets arriving subsequently to the receipt of the acknowledgment packet are designated to the next burst.

It takes a finite propagation time across the network $t_{\mathrm{prop,net}}$ for the first bit to arrive at the destination edge router, so that the reserved wavelength is idle and not used for data transmission for the period $t_{\mathrm{idle}} = t_{\mathrm{prop,ack}} + t_{\mathrm{prop,net}}$. The time to complete the burst transmission is $t_{\mathrm{trans}} = L_{\mathrm{burst}}/b_{\mathrm{core}}$, so that the wavelength holding time is given by $t_{\mathrm{WHT}} = t_{\mathrm{idle}} + t_{\mathrm{trans}}$. In principle, the wavelength holding time could be fixed either by the maximum edge delay or by streaming data, in which case t_{WHT} would be less predictable but the lightpath utilization would increase [11]. The maximum deterministic latency or upper bound on the maximum transmission time that packets experience between entering the core network at the source and leaving the destination routers is

$$\mathrm{Latency}_{\max} = t_{\mathrm{edge}} + t_{\mathrm{prop,net}} + \frac{L_{\mathrm{burst}}}{b_{\mathrm{core}}}. \tag{1}$$

The arrival of the acknowledgment packet from the control node sets the start of the subsequent burst assembly and cycle repeats.

From the analysis of the timings involved in burst assembly and transmission it is clear that the network efficiency depends on the processing speed of the network controller. Minimization of t_{proc} can be achieved by applying fast dynamic routing and wavelength assignment algorithms. Efficient algorithms already exist for the optimization of static and dynamic WRONs; see, for example [1], [17]. Because the focus of this paper is on the analysis of the effects of traffic statistics on the OBS network and the evaluation of an upper bound to the performance of any RWA algorithm, it was assumed in this work that a wavelength will always be available, under the conditions of an ideal RWA algorithm. For a given network topology and optimized route look up and wavelength allocation algorithms $t_{\mathrm{prop,ack}}$ and t_{proc} are known *a priori* and the timings of wavelength requests can be adjusted by the edge routers to meet latency and PLR criteria.

C. Modeling of the Impact of Traffic Statistics on Burst Aggregation

The traffic statistics of the arriving packets is likely to have a significant impact on burst aggregation, buffering, and the resultant network performance and, most important, the edge delay t_{edge} and the PLR. To analyze this, the following simulations were carried out using a single first in–first out (FIFO) queue. The incoming traffic was generated using an ON–OFF source at the input of the edge router with independent probability density functions (pdfs) for the ON state, P(ON), and the OFF state, P(OFF), to allow variation of both packet length and packet interarrival time. Telephone call arrivals and call holding were modeled by Poisson interarrival time and exponential call holding times, but this model may not hold for the description of data traffic [18], although the correct model of data traffic is under much debate and depends on the implemented protocol (such as IP or Ethernet) [19], [20]. For bursty traffic and finite values of the edge delays considered in this work, a possible traffic model is realized by the multiplexing of several heavy-tailed Pareto distributions, given by $P(t) = (\alpha \cdot A^{\alpha})/t^{\alpha+1}$, where $1 < \alpha < 2$, $A > 0$ and $t \geq A$. Simulations were carried out for different scenarios to calculate the distribution of the burst size L_{burst} and the resulting PLR for a finite length buffer [5]. The pdfs applied for traffic modeling included Pareto, Poisson, and fixed packet length and packet interarrival time distributions. A minimum packet length of 50 bytes, approximately the size of a short IP packet (40 bytes IPv4, 60 bytes IPv6) or an asynchronous transfer mode (ATM) cell (53 bytes) was assumed, in combination with a buffer size $B =$ 400 Mb (47.7 MB) for an average input bit rate $b_{\mathrm{in}} = 10$ Gb/s into a single buffer with uniformly distributed destination addresses, as explained in Section II-A.

To reduce the required header processing, future networks might operate with packets which are significantly longer than minimum IP packet size. The decreased granularity or *fragmentation* will typically be determined by the network protocol and the optimum level of fragmentation requires analysis. Here, different levels of packet fragmentation were modeled and the results shown for values ranging from 50 bytes to 5 kB. The low value corresponds to the current data networks, in which 40-byte TCP–IP acknowledgments account for more than 50% of the total traffic. Longer packets, however, may simplify the processing and forwarding functions and future applications for data transfer or multimedia applications may make use of longer IP packets that map minimum packet lengths up to 5 kB.

III. Edge Router Simulation Results

Fig. 4 shows the resultant burst size distribution as a function of t_{edge} and resulting PLR for a minimum packet size of 5 kB with $b_{\mathrm{in}} = 10$ Gb/s and an average load of 0.1 (i.e., max. access buffer bandwidth 100 Gb/s) for the following packet and interarrival time distributions:

Case 1) Pareto ($\alpha = 1.5$) packet length distribution, Pareto ($\alpha = 1.5$) interarrival time distribution;

Case 2) fixed length packet sizes, Pareto interarrival time distribution ($\alpha = 1.5$);

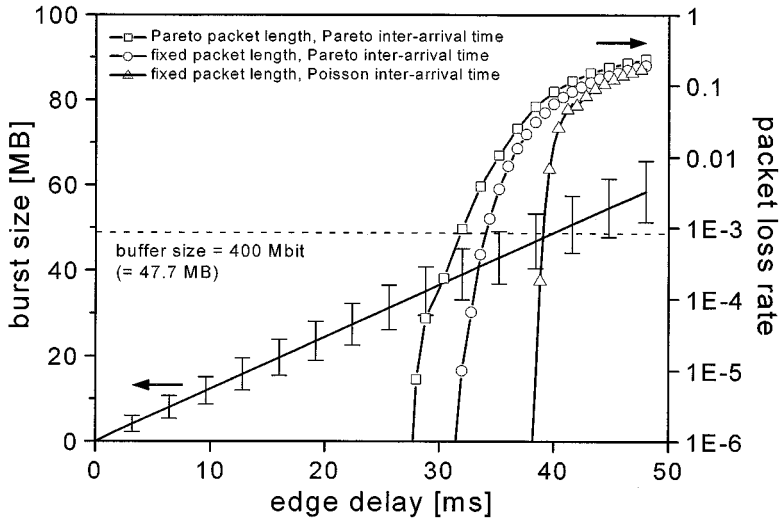


Fig. 4. Simulation results for burst size and PLR as a function of t_{edge} and a mean input bit rate $b_{\text{in}} = 10$ Gb/s. The burst size increase is plotted for an infinitely large buffer size, but PLR for a finite buffer size $B = 400$ Mbit (47.7 MB). Bars indicate 95% confidence level for Pareto–Pareto distribution with minimum packet size of 5 kB.

Case 3) fixed length packet sizes, Poisson interarrival time distribution.

In all three cases, it can be seen that the mean burst size increases linearly. However, the burst size distribution for a given t_{edge} does not follow the same behavior. For burst statistics on finite time scales (as for a given t_{edge}), a possible measure of the burstiness is the variance of the burst size distribution. Packet loss variation according to the burstiness of the input traffic for cases 1) to 3) is shown by the PLR curves in Fig. 4. Because the core network is free of blocking and, therefore, no packet loss in the core was assumed throughout this paper, packet loss refers to those packets lost due to buffer overflow in edge routers. The largest deviation of the burst size distribution for a given t_{edge} was observed for case 1), indicated by bars for a distribution with 95% confidence level. For the calculation of the burst size distribution, an infinite buffer size was assumed. For the calculation of the PLR, the buffer size then was bounded to $B = 400$ Mb (47.7 MB). For finite simulation time, an average PLR of 10^{-6} was reached for edge delays of 27.5, 31.5, and 38 ms for cases 1)–3). The results can be compared to the case of a continuous bit rate (CBR), also referred to as the fluid traffic model, with an achievable edge delay of 40 ms before packet loss occurs. The application of a CBR traffic model allows the development of an analytical model independent from the actual traffic statistics, which can be applied to derive bounds for parameters. The PLR graphs in Fig. 4 show that bursty traffic significantly reduces the maximum allowable t_{edge}. To meet a specific PLR, e.g., 10^{-6}, the maximum allowable t_{edge} before releasing a burst would be < 28 ms. This is important for all applications and network services whose quality is determined by a low PLR, such as voice transmission. It is emphasized that the value of 40 ms in this example is the upper limit before packet loss occurs; for time critical applications, the buffer can be emptied at a faster rate, although best effort type traffic would experience longer edge delays, after which there is no further delay except the propagation time.

More detailed analysis of the variation of the PLR for case 1) is shown in Fig. 5. Increase in PLR is observed for $t_{\text{edge}} = 28.8$ ms. Fig. 6 (plotted for the same values) shows the probability P for the PLR to exceed a given threshold X for t_{edge} varying in the range from 28.8 to 41.6 ms. A comparison with the average PLR in Fig. 4 shows that for an edge delay of 33.6 ms, for which an average PLR of $3.83 \cdot 10^{-3}$ was calculated, PLR > 0.08 appears with a probability of 1%. The results signify that the variation in the PLR must be taken into account to accurately characterize the QoS of a lightpath.

The effects of packet fragmentation on the PLR and maximum allowable edge delay were also analyzed by using the same statistics as in case 1), but with a minimum packet length of 5 kB. The resultant PLRs are shown in Fig. 7 and result in maximum allowable t_{edge} of 27.5, 34, and 36 ms, respectively, achieving mean PLR $< 10^{-6}$.

The same figure shows that, for aggregation of packets over timescales significantly longer than the packet length, the burst size distribution can be approximated by a normal distribution. For the assumed buffer size of 400 Mb (47.7 MB), the PLR values derived from a Gaussian approximation are in good agreement with those obtained from simulation for minimum packet sizes ≤ 5 kB. The result that the burst size distribution is normal is important because it proves that the central limit theorem can be applied, simplifying the modeling of the burst aggregation process over timescales. This simplifies the analysis and enables the scaling of the mean and the variance of the burst size with the edge delay, for the basic stochastic processes for the packet length and packet interarrival time. With the burst size distribution approaching a normal distribution with a set of mean and variance $\{\mu, \sigma^2\}$, the PLR can then be explicitly calculated using the error function for a given buffer size B.

IV. Core Network Performance Theory and Results

In this section, an analytical model to calculate network performance parameters is derived for some of the simulation

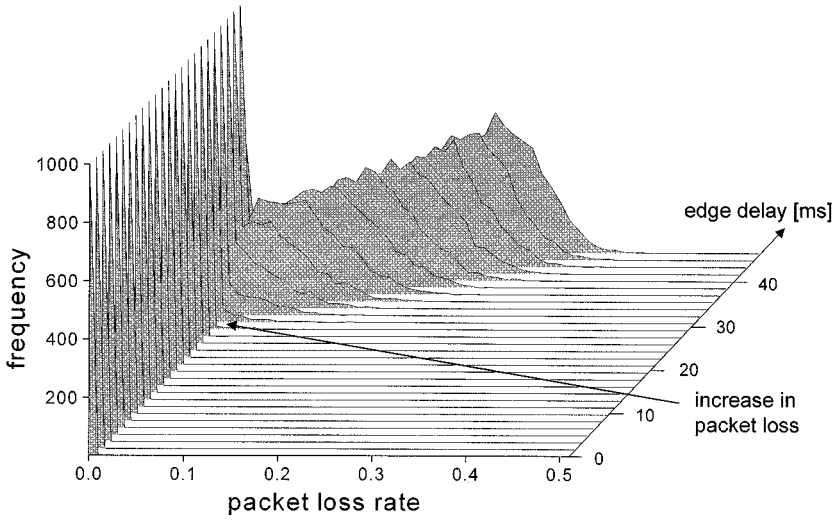


Fig. 5. Deviations of the PLR for Pareto packet, Pareto interarrival time statistics ($\alpha = 1.5$) for $B = 400$ Mbit (47.7 MB) and minimum packet size of 5 kB. For clarity, frequency values > 1000 were omitted.

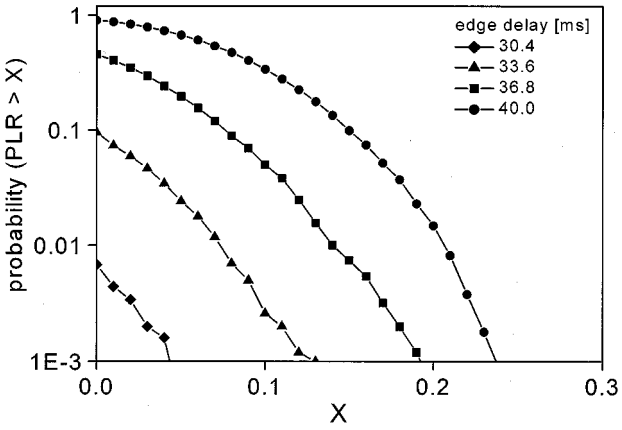


Fig. 6. Probability P for PLR $> X$ (cumulative distribution) for Pareto packet size ($\alpha = 1.5$), Pareto interarrival time distribution ($\alpha = 1.5$), minimum packet size of 5 kB.

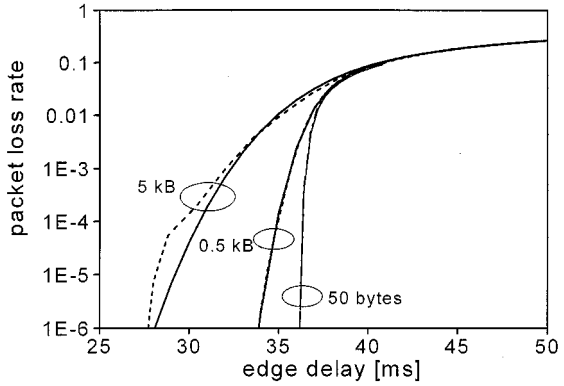


Fig. 7. Simulation results for the PLR as a function of the edge delay for $B = 400$ Mb (47.7 MB) and a mean input bit rate $b_{\text{in}} = 10$ Gb/s for different levels of packet size, 5 kB (dash), 0.5 kB (dot), and 50 bytes (dash-dot). The PLR calculated from burst size distribution (assumed Gaussian) is shown by a solid line.

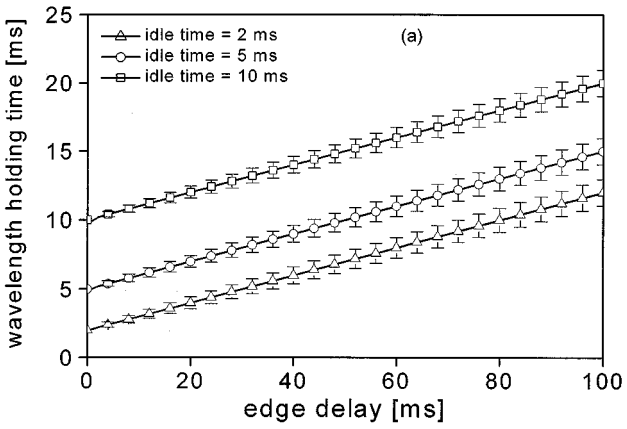
results of the previous section to extend the analysis of burst aggregation and to quantify the achievable wavelength reuse and lightpath utilization in OBS networks for this fast circuit-switched architecture where the lightpaths are set up only for the time required to transmit the content of a single buffer between two edge routers and then released for subsequent requests. This time includes an overhead required for lightpath setup and propagation delays; we refer to it as idle time t_{idle} in the remainder of the paper. Following the results of the previous section (see Fig. 4), in this section, it is assumed that burst sizes increase linearly, equivalent to the case of CBR traffic arriving to the buffer and for which there is no variation in the burst size. Then, for a constant load and CBR traffic, the burst size L_{burst} is proportional to the edge delay and the input bit-rate b_{in}, so that $L_{\text{burst}} = b_{\text{in}} \cdot t_{\text{edge}}$.

For clarity and simplicity, the analysis in this section is based only on mean values for all parameters. However, for an arbitrary burst size distribution, the pdf can be derived as described in Appendix A.

Once a burst is assigned to a free wavelength, this wavelength will be reserved and is used until the buffer content is transmitted from source to the destination edge router. The wavelength holding time t_{WHT}, shown in Fig. 3, can be thought of as equivalent to the call-holding time in circuit-switched networks. It is given by

$$t_{\text{WHT}} = t_{\text{idle}} + \frac{L_{\text{burst}}}{b_{\text{core}}} = t_{\text{idle}} + \frac{1}{A} \cdot t_{\text{edge}} \tag{2}$$

where t_{idle} is the idle time before the burst reaches the destination edge router plus the time for the acknowledgment and A is the core bit-rate to input bit-rate ratio $A = b_{\text{core}}/b_{\text{in}}$. For small values of A, the data transmission time t_{trans} can be in the range of tens of milliseconds, so that $t_{\text{idle}} \ll t_{\text{WHT}}$. Time t_{idle} starts to affect the service quality when the values of the t_{trans} are comparable to t_{idle} and dominate the wavelength holding time for high core bit rates such as $b_{\text{core}} = 100$ Gb/s, as shown in Fig. 8(a) for $t_{\text{idle}} = 2, 5, 10$ ms. Fig. 8(b) shows the effect for $t_{\text{idle}} = 5$ ms and a variation of b_{core} from 20 to 100 Gb/s, where, for $b_{\text{core}} = 20$ Gb/s, the t_{WHT} is significantly longer than for $b_{\text{core}} = 100$ Gb/s.



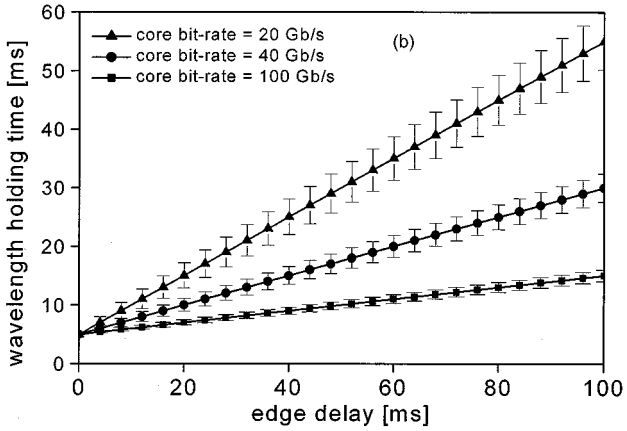
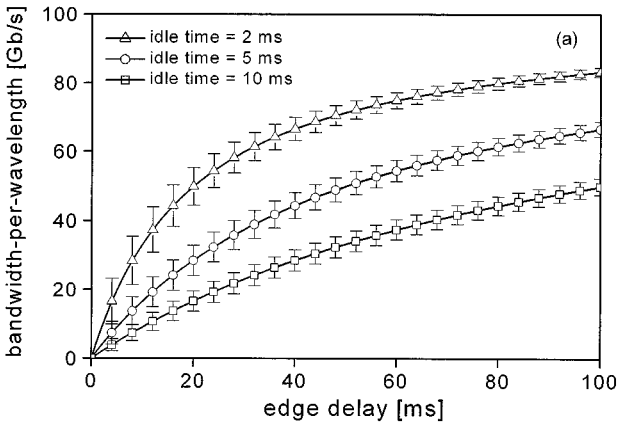


Fig. 8. Wavelength holding time t_{WHT} as a function of the edge delay for (a) $b_{\mathrm{core}} = 100$ Gb/s, $b_{\mathrm{in}} = 10$ Gb/s and $t_{\mathrm{idle}} = 2, 5, 10$ ms and $t_{\mathrm{idle}} = 5$ ms and (b) $b_{\mathrm{core}} = 20, 40, 100$ Gb/s. Bars show 95% confidence level.



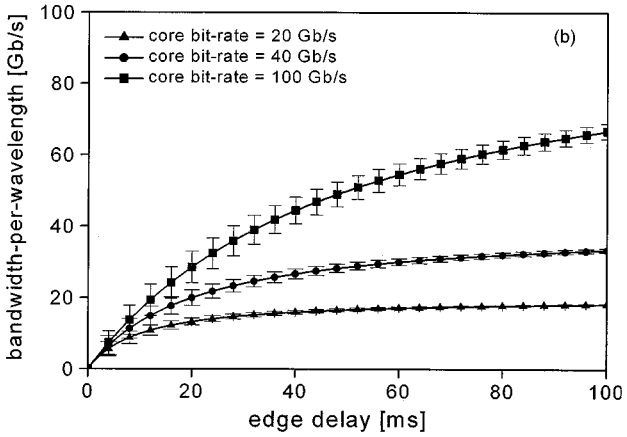


Fig. 9. Bandwidth per wavelength $B_{\mathrm{per}\lambda}$ for (a) $b_{\mathrm{core}} = 100$ Gb/s, $b_{\mathrm{in}} = 10$ Gb/s, and $t_{\mathrm{idle}} = 2, 5, 10$ ms and for (b) $t_{\mathrm{idle}} = 5$ ms and $b_{\mathrm{core}} = 20, 40, 100$ Gb/s. Bars show 95% confidence level.

A parameter following from (2) is the bandwidth per wavelength $B_{\mathrm{per}\lambda}$, which indicates the effective bandwidth of a lightpath used for transmission of data between edge routers

$$B_{\mathrm{per}\lambda} = \frac{L_{\mathrm{burst}}}{t_{\mathrm{WHT}}} = \frac{b_{\mathrm{in}} \cdot t_{\mathrm{edge}}}{t_{\mathrm{idle}} + \frac{1}{A} \cdot t_{\mathrm{edge}}}. \tag{3}$$

The influence of t_{idle} on $B_{\mathrm{per}\lambda}$ is shown in Fig. 9(a) for $b_{\mathrm{core}} = 100$ Gb/s and $t_{\mathrm{idle}} = 2, 5, 10$ ms. The increase in bandwidth for the identical values t_{edge} is reduced for higher t_{idle}; for $t_{\mathrm{idle}} = 10$ ms, values remain below 50 Gb/s for 100-Gb/s physical bit rate. Fig. 9(b) shows the effect of bandwidth saturation for $t_{\mathrm{idle}} = 5$ ms and core bit rates varying from 20 to 100 Gb/s. The significance of the results is that $B_{\mathrm{per}\lambda}$ remains significantly smaller than the optical line rate for $t_{\mathrm{edge}} \leq 40$ ms, especially for high b_{core}, such as 100 Gb/s.

Relating the bandwidth per wavelength $B_{\mathrm{per}\lambda}$ to the physical bit rate in the core b_{core} leads to the dimensionless parameter U, the utilization that describes the efficiency with which the lightpath bandwidth is utilized

$$U = \frac{B_{\mathrm{per}\lambda}}{b_{\mathrm{core}}} = \frac{t_{\mathrm{edge}}}{A \cdot t_{\mathrm{idle}} + t_{\mathrm{edge}}}. \tag{4}$$

Maximizing the use of available resources is key for the network operator, implying that utilization U must be maximized. Fig. 10 shows the dependence of U on t_{edge}, A, and t_{idle} for 0 ms $\leq t_{\mathrm{edge}} \leq 200$ ms, $0 \leq A \leq 100$ and for $t_{\mathrm{idle}} = 2$ [see Fig. 10(a)] and 10 ms [see Fig. 10(b)], as calculated in the previous section. As can be seen from Fig. 10, the highest bandwidth utilization is achieved for low values of A and high edge delays (> 50 ms). Utilization also increases for smaller t_{idle} but the same values of t_{edge} and A, as shown for $t_{\mathrm{idle}} = 2$ ms in Fig. 10(a), as compared to $t_{\mathrm{idle}} = 10$ ms in Fig. 10(b).

In high-speed networks, it can be assumed that $b_{\mathrm{core}} \gg b_{\mathrm{in}}$ results in $t_{\mathrm{WHT}} \ll t_{\mathrm{edge}}$, i.e., the time required to aggregate a burst is significantly larger than the time to transmit it. In the case of dynamic wavelength allocation, an unused wavelength can be assigned to another edge router and the resultant increase in the wavelength reuse can be defined as a wavelength reuse factor (RUF), defined as

$$\mathrm{RUF} = \frac{t_{\mathrm{edge}}}{t_{\mathrm{WHT}}} = \frac{A \cdot t_{\mathrm{edge}}}{A \cdot t_{\mathrm{idle}} + t_{\mathrm{edge}}} = A \cdot U. \tag{5}$$

For consistency, the variation of RUF is plotted in Fig. 11 for the same range of values as for the utilization U. For comparison to a static WRON, Fig. 11 shows the values for RUF $= 1$. This is justified by the assumption that in a static WRON, a given lightpath is established for a long period, but not shared between different edge routers. In an optical network with dynamic wavelength assignment, this is equivalent to a lightpath permanently assigned between two edge routers, i.e., $t_{\mathrm{edge}} =$

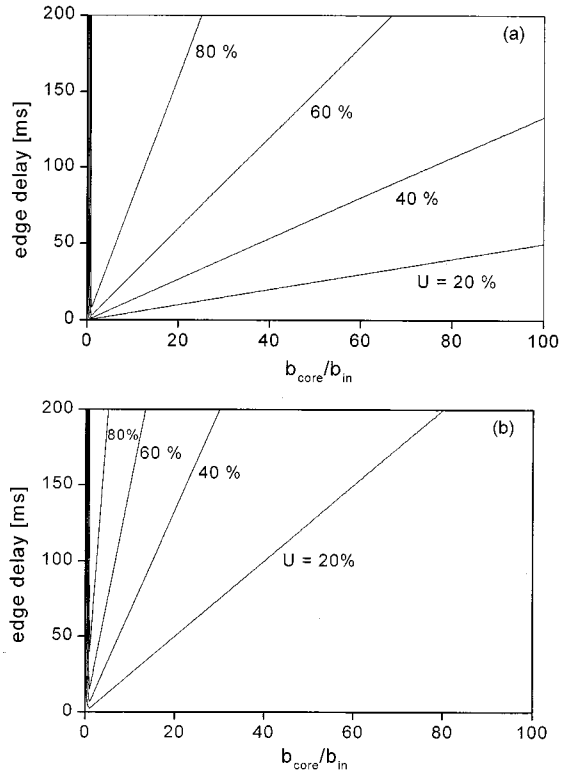


Fig. 10. Mean bandwidth utilization U as function of the edge delay $t_{\rm edge}$ and bit-rate ratio A for (a) $t_{\rm idle} = 2$ ms and (b) $t_{\rm idle} = 10$ ms.

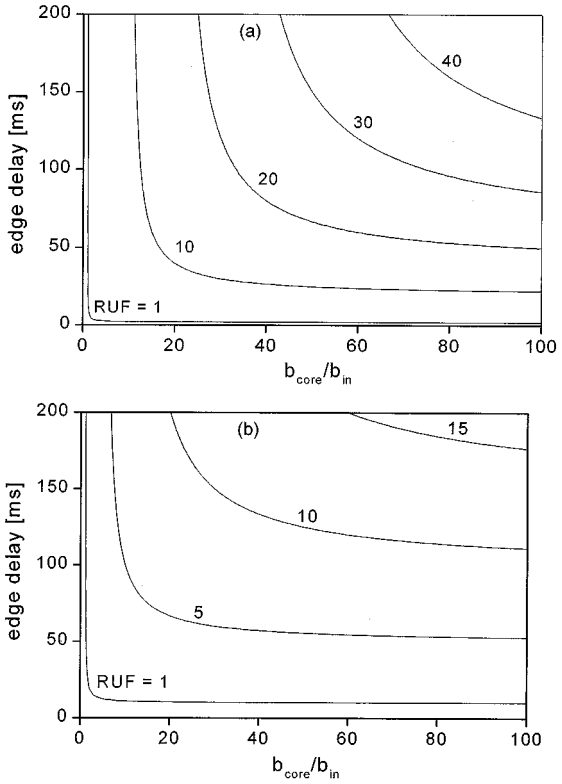


Fig. 11. Mean wavelength RUF as function of edge delay $t_{\rm edge}$ and bit-rate ratio A for (a) $t_{\rm idle} = 2$ ms and (b) $t_{\rm idle} = 10$ ms.

$t_{\rm WHT}$. For RUF < 1, the WR-OBS network would theoretically require more wavelengths than in a static WRON to satisfy all demanded connections and, therefore, values for RUF < 1 represent the region of network instability where the total input load exceeds the network throughput.

Despite the potential savings in terms of the number of wavelengths, it should be noted that the actual number of wavelengths required also depends on the physical topology and routing strategy, as well as the wavelength allocation algorithm [1], [17].

The variation of the mean RUF for $t_{\rm idle} = 2$ and 10 ms are shown in Fig. 11(a) and (b). It can be seen that the RUF increases with both $t_{\rm edge}$ and A, to values of $\text{RUF}_{\max} = 50$ and 16.7, respectively, for the given range of $t_{\rm edge}$ and A, with $A_{\max}$, $t_{\rm edge,max}$ as the maximum values of their given ranges ($A_{\max} = 100$, $t_{\rm edge,max} = 200$ ms). It should be noted that for a given $t_{\rm idle}$, upper bounds for the reuse factor can be determined for either constant $t_{\rm edge}$ or constant A.

In the case of $t_{\rm edge} =$ const. and increasing A, the data transmission time $L_{\rm burst}/b_{\rm core}$ becomes negligible so that $t_{\rm WHT} \approx t_{\rm idle}$ and RUF $\leq t_{\rm edge}/t_{\rm idle}$. In the case of $A =$ const., the buffer content increases proportionally with $t_{\rm edge}$, restricting the reuse factor to RUF $\leq A$. For constant $t_{\rm edge}$, an increase of A is only useful for $A \leq t_{\rm edge}/t_{\rm idle}$; for larger values of A, the reuse factor will only increase marginally. For constant A, an increase in $t_{\rm edge}$ is beneficial only for $t_{\rm edge} \leq A \cdot t_{\rm idle}$.

A comparison between Figs. 10 and 11 shows that both utilization and reuse factor increase with $t_{\rm edge}$, but that U is maximum for low values of A, whereas RUF maximizes for high values of A. The resulting tradeoff for constant $t_{\rm edge}$ between both parameters can be described by the dimensionless parameter P defined as the product of U and RUF

$$P = U \cdot \text{RUF} \tag{6}$$

and is plotted in Fig. 12 for 0 ms $\leq t_{\rm edge} \leq$ 200 ms, $0 \leq A \leq 100$, with $t_{\rm idle} = 2$ ms [see Fig. 12(a)] and 10 ms [see Fig. 12(b)]. For a constant value of $t_{\rm edge}$, P can be optimized for a set of parameters $\{A, t_{\rm edge}\}$ such that $U_{\rm opt} = 50\%$ and $\text{RUF}_{\rm opt} = 2 \cdot P_{\rm opt}$ are achievable. Hence, the OBS network benefits from good utilization and wavelength reuse. As in the preceding graphs, the optimization process also depends on $t_{\rm idle}$, as shown in Fig. 12(a) and (b). Especially for $t_{\rm idle} = 10$ ms, RUF ≥ 1 must be maintained to avoid network instability.

To investigate the impact of the $t_{\rm idle}$ on the network, both the utilization U and reuse factor RUF are replotted for different values of edge delays (10, 20, 50 ms) and constant $A = 10$ in Figs. 13 and 14. To ensure that RUF > 1 requires $t_{\rm idle} < (A-1)/A \cdot t_{\rm edge} \Leftrightarrow t_{\rm idle} < t_{\rm edge}$ for $A \gg 1$. A key result is that, for $A \gg 1$, as in high core bit-rate networks, a high reuse factor can be achieved only for $t_{\rm idle}$ on the timescale of a few milliseconds. It is important to note that, in order to achieve efficient wavelength reuse, the lightpath setup time must be as small as possible and for a fixed $t_{\rm edge}$, $\text{RUF}_{\max}$ is given for instantaneous lightpath setup ($t_{\rm idle} = 0$) as $\text{RUF}_{\max} = A$. Not only does the wavelength reuse factor decrease with an increasing $t_{\rm idle}$, but so does the lightpath utilization, which in all cases is less than 50% for idle times $t_{\rm idle} \geq 10$ ms and drops sharply especially for an edge delay of 10 ms. These results show that the idle time $t_{\rm idle}$ is a key parameter in the design of OBS networks

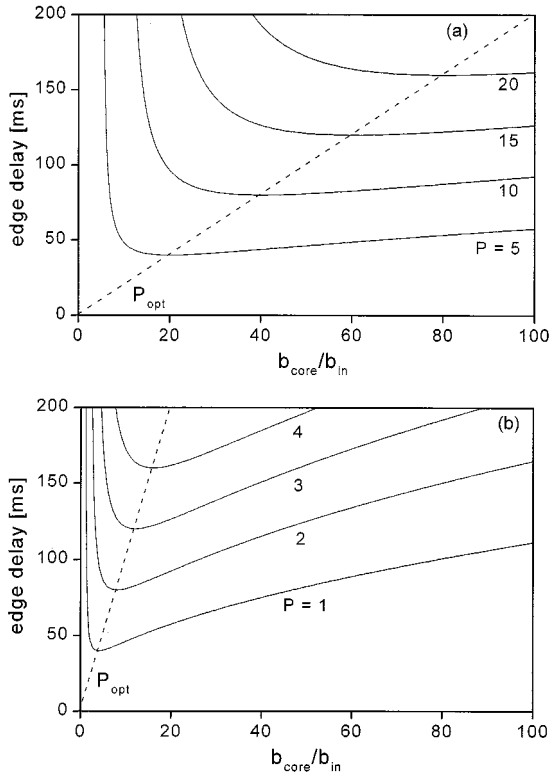


Fig. 12. Network performance parameter P showing the tradeoff between utilization U and wavelength reuse RUF for constant $t_{\rm edge}$, (a) $t_{\rm idle} = 2$ ms and (b) $t_{\rm idle} = 10$ ms. Optimum values $P_{\rm opt}$ are shown by the dashed line.

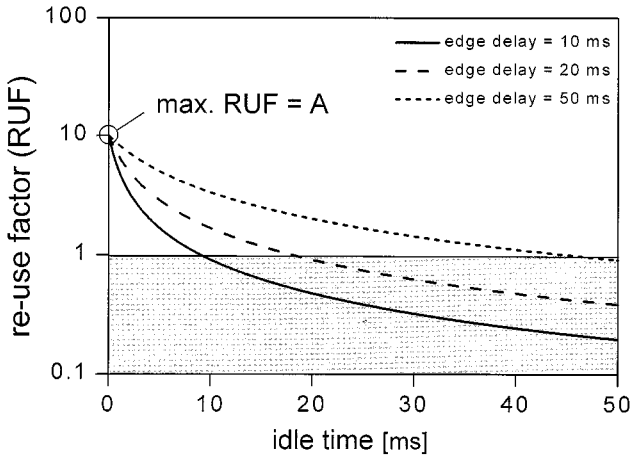


Fig. 13. Wavelength RUF as a function of the idle time ($t_{\rm idle}$) for $t_{\rm edge} = 10$, 20, 50 ms, and bit-rate ratio $A = 10$. Shaded region: network requires more wavelengths than in a static WRON.

with dynamic wavelength allocation and define the performance requirements on the dynamic RWA algorithm used for lightpath establishment between edge routers to minimize the overhead of the time $t_{\rm idle}$ to achieve the operational advantage of increased throughput per wavelength under dynamic wavelength operation. For the speed of a RWA algorithm, this implies that the RWA decision time must not exceed the edge delay and is constrained $\leq$ 10 ms to meet even tight delay constraints. Identifying these constraints will help optimize the RWA process carried out in the control node to maximize the number of edge routers and network routes.

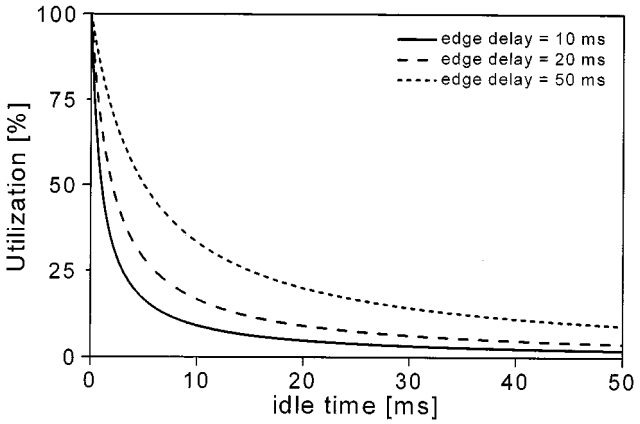


Fig. 14. Lightpath utilization U as a function of the idle time ($t_{\rm idle}$) for $t_{\rm edge} = 10, 20, 50$ ms, and bit-rate ratio $A = 10$.

From the analysis carried out, it is clear that the constraints for $t_{\rm idle}$ limit the network diameter or the allowable minimum edge delays for efficient network operation. Thus, a WR-OBS scheme brings most advantages for network sizes found in Europe or metropolitan area-type networks where the lower signaling round-trip times allow wavelength savings from dynamic network operation.

V. Conclusion

This paper describes and analyzes a WR-OBS network that combines the functions of OBS with fast circuit switching by dynamically assigning and releasing wavelength-routed lightpaths over a bufferless optical core. The potential advantages of this architecture compared to conventional OBS are explicit QoS provisioning and, compared to static WRONs, are in fast adaptation to dynamic traffic changes in optical networks and more efficient utilization of each wavelength channel. The proposed architecture ensures a deterministic delay for the optical packets through a known, predefined queueing delay at the edge and burst aggregation and the propagation in the core network. Moreover, it guarantees an acknowledgment of the wavelength assignment for QoS-determined provisioning and uses dynamic wavelength routing. The achievable edge delays were calculated for different traffic statistics and it was shown that the allowable edge delay to maintain a predefined mean PLR was significantly reduced in the presence of bursty traffic, resulting in more frequent wavelength requests and less efficient use of buffer resources. Similar degradation of buffer performance was observed for different levels of allowable minimum packet size or *fragmentation*, varied between 50 bytes and 5 kB. Simulations also showed that the burst size distribution could be approximated by a Gaussian function, hence allowing the development of an analytical description and model of the network.

The lightpath utilization and wavelength reuse factor were introduced to characterize OBS networks with dynamic wavelength assignment. It was shown that these parameters could be described as a function of the edge delay $t_{\rm edge}$, the idle time $t_{\rm idle}$, and the ratio of core to input bit rates A, allowing the results to be generalized to cover a wide range of input and core bandwidths. These results also allowed to quantify the operating

range for A and t_{edge} for which increases in lightpath utilization and reuse of a given wavelength channel and thus increased network throughput can be achieved relative to static WRONs. However, to attain this performance, the signaling round-trip time for the acknowledgment of dynamic wavelength reservation and wavelength assignment must be much shorter than the edge delay, setting stringent limits on the performance of such networks. The allowable round-trip time delay is related to the maximum network diameter. Whereas the WR-OBS architecture would probably work for maximum network diameters up to 6000-km long with round-trip time delays of 30 ms, it would offer the most significant advantage for networks with smaller diameters, e.g., as found with European network operators and for metropolitan-area size networks. The results can be applied to the design and the dimensioning of wavelength-routed optical burst-switched networks and the optimization of scheduling, control, and wavelength assignment in coordination between the electronic and the optical network layers.

APPENDIX

In this appendix, the deviation of the network performance parameters t_{WHT}, $B_{\mathrm{per}\lambda}$, U, and RUF from a given pdf for the burst size distribution $f_{L_{\mathrm{burst}}}$ is shown using probability theory [21]. For a given random variable (RV) X with pdf f_X, the pdf f_Y of a new RV Y with $Y = g(X)$ can be described by the following equation:

$$f_Y(Y) = \frac{f_X(X)}{\left|\frac{\partial g}{\partial X}\right|} \tag{A1}$$

assuming $X = g^{-1}(Y)$ exists and $\partial g/\partial X \neq 0$. The pdfs of the network performance parameters, defined in Section IV, are derived assuming that $X = L_{\mathrm{burst}}$. Hence, L_{burst} must be expressed as a function of the respective performance parameter for calculation of the pdf. Based on (A1), the burst size and the pdf f_{WHT} are derived as a function of the wavelength holding time as follows:

$$L_{\mathrm{burst}}\left(t_{\mathrm{WHT}}\right) = \left(t_{\mathrm{WHT}} - t_{\mathrm{idle}}\right) \cdot b_{\mathrm{core}} \tag{A2}$$

$$f_{\mathrm{WHT}} = f_{L_{\mathrm{burst}}}\left[L_{\mathrm{burst}}\left(t_{\mathrm{WHT}}\right)\right] \cdot b_{\mathrm{core}}. \tag{A3}$$

The variations of t_{WHT} were indicated in Fig. 8 by bars for a 95% confidence level. The pdf of the bandwidth per wavelength is

$$L_{\mathrm{burst}}\left(B_{\mathrm{per}\lambda}\right) = \frac{B_{\mathrm{per}\lambda} \cdot t_{\mathrm{idle}}}{1 - \frac{B_{\mathrm{per}\lambda}}{b_{\mathrm{core}}}} \tag{A4}$$

$$f_{B_{\mathrm{per}\lambda}} = f_{L_{\mathrm{burst}}}\left[L_{\mathrm{burst}}\left(B_{\mathrm{per}\lambda}\right)\right] \cdot \frac{\left(t_{\mathrm{idle}} + \frac{L_{\mathrm{burst}}\left(B_{\mathrm{per}\lambda}\right)}{b_{\mathrm{core}}}\right)^2}{t_{\mathrm{idle}}}. \tag{A5}$$

In Fig. 9, the variations of $B_{\mathrm{per}\lambda}$ were indicated by bars for a 95% confidence interval. The results for the respective pdf of the utilization are as follows:

$$L_{\mathrm{burst}}\left(U\right) = \frac{U \cdot b_{\mathrm{core}} \cdot t_{\mathrm{idle}}}{1 - U} \tag{A6}$$

$$f_U = f_{L_{\mathrm{burst}}}\left[L_{\mathrm{burst}}\left(U\right)\right] \cdot \frac{\left(t_{\mathrm{idle}} + \frac{L_{\mathrm{burst}}\left(U\right)}{b_{\mathrm{core}}}\right)^2}{t_{\mathrm{idle}}} \cdot b_{\mathrm{core}}. \tag{A7}$$

The pdf of the reuse factor is determined as follows:

$$L_{\mathrm{burst}}\left(\mathrm{RUF}\right) = \frac{t_{\mathrm{edge}} - \mathrm{RUF} \cdot t_{\mathrm{idle}}}{\mathrm{RUF}} \cdot b_{\mathrm{core}} \tag{A8}$$

$$f_{\mathrm{RUF}} = f_{L_{\mathrm{burst}}}\left[L_{\mathrm{burst}}(\mathrm{RUF})\right] \cdot \left(t_{\mathrm{idle}} + \frac{L_{\mathrm{burst}}(\mathrm{RUF})}{b_{\mathrm{core}}}\right)^2 \cdot \frac{b_{\mathrm{core}}}{t_{\mathrm{edge}}}. \tag{A9}$$

ACKNOWLEDGMENT

The authors would like to thank IEEE LEOS for the award of the Graduate Student Fellowship in 2001 to M. Düser; J. E. Midwinter, G. Gavioli, R. I. Killey, E. Kozlovski, A. Myers (all UCL), and D. Wischik of the Statistical Laboratory, Cambridge University, Cambridge, U.K.; S. P. Ferguson of Marconi Communications and I. De Miguel of the University of Valladolid, Valladolid, Spain, for invaluable discussions and suggestions throughout this work; and the referees for detailed critical reading of the manuscript and suggestions that helped clarify and improve this paper.

REFERENCES

[1] S. Baroni, P. Bayvel, R. J. Gibbens, and S. K. Korotky, "Analysis and design of resilient multifiber wavelength-routed optical transport networks," *J. Lightwave Technol.*, vol. 17, pp. 743–758, May 1999.

[2] C. Qiao and M. Yoo, "Choices, features and issues in optical burst switching," *SPIE Opt. Networks Mag.*, vol. 1, no. 2, pp. 36–44, April 2000.

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

[4] K. Dolzer, C. Gauger, J. Späth, and S. Bodamer, "Evaluation of reservation mechanisms for optical burst switching," *AEÜ Int. J. Electron. Commun.*, vol. 55, no. 1, pp. 1–8, Jan. 2001.

[5] M. Düser, E. Kozlovski, R. I. Killey, and P. Bayvel, "Design trade-offs in optical burst switched networks with dynamic wavelength allocation," in *Proc. Eur. Conf. Optical Communication*, vol. 2, Munich, Germany, Sept. 2000, pp. 23–24.

[6] C. Qiao, "Labeled optical burst switching for IP-over-WDM integration," *IEEE Commun. Mag.*, vol. 38, pp. 104–114, Sept. 2000.

[7] J. Y. Wei, J. L. Pastor, R. S. Ramamurthy, and Y. Tsai, "Just-in-time optical burst switching for multiwavelength networks," in *Proc. IFIP Conf. Broadband '99*, Hong Kong, China, Nov. 1999, pp. 339–352.

[8] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP-over-WDM networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2062–2071, Oct. 2000.

[9] Y. Xiong, M. Vandenhoute, and H. C. Cankaya, "Control architecture in optical burst-switched networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1838–1851, Oct. 2000.

[10] S. Verma, H. Chaskar, and R. Ravikanth, "Optical burst switching: A viable solution for terabit IP backbone," *IEEE Network*, vol. 14, no. 6, pp. 48–53, Nov./Dec. 2000.

[11] I. de Miguel, M. Düser, and P. Bayvel, "Traffic load bounds for optical burst-switched networks with dynamic wavelength allocation," in *Toward an Optical Internet*, A. Jukan, Ed. Norwell, MA: Kluwer, 2002, vol. 1, pp. 209–224.

[12] A. Bianco, E. Leonardi, M. Mellia, and F. Neri, "Network controller design for SONATA—A large-scale all-optical passive network," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2017–2028, Oct. 2000.

[13] S.-T. Chuang, A. Goel, N. McKeown, and B. Prabhakar, "Matching output queueing with a combined input/output-queued switch," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1030–1039, June 1999.

[14] C.-K. Chan, K. L. Sherman, and M. Zirngibl, "A fast 100-channel wavelength-tuneable transmitter for optical packet switching," *IEEE Photon. Technol. Lett.*, vol. 13, pp. 729–731, July 2001.

[15] E. Kozlovski and P. Bayvel, "QoS performance of WR-OBS network architecture with request scheduling," in *Proc. IFIP 6th Working Conf. Optical Network Design Modeling* , Torino, Italy, Feb. 4–6, 2002, pp. 101–116.

[16] J. Gripp, M. Duelk, J. Simsarian, S. Chandrasekhar, P. Bernasconi, and A. Bhardwaj *et al.*, "Demonstration of a 1.2 Tb/s optical packet switch fabric ($32*40$ Gb/s) based on 40 Gb/s burst-mode clock-data-recovery, fast tuneable lasers and a high-performance NxN AWG," in *Proc. 27th European Conf. Optical Commun.*, vol. 6, Amsterdam, The Netherlands, Oct. 2001, pp. 58–59.

[17] A. Mokhtar and M. Azizoglu, "Adaptive wavelength routing in all-optical networks," *IEEE/ACM Trans. Networking*, vol. 6, pp. 197–206, Apr. 1998.

[18] V. Paxson and S. Floyd, "Wide area traffic: The failure of poisson modeling," *IEEE/ACM Trans. Networking*, vol. 3, pp. 226–244, June 1995.

[19] L. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of ethernet traffic (extended version)," *IEEE/ACM Trans. Networking*, vol. 2, pp. 1–15, Feb. 1994.

[20] W. Willinger, "Self-similarity in wide-area network traffic," in *Proc. IEEE Conf. Lasers Electro-Optics*, vol. 2, 1997, pp. 462–463.

[21] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991, pp. 87–102.

Michael Düser (S'99) received the Dipl.-Ing. degree with distinction in electrical engineering from the University of Dortmund, Dortmund, Germany, in 1998. He is currently pursuing the Ph.D. degree in the Optical Networks Group, Department of Electronic and Electrical Engineering, University College London, London, U.K.

From November 1997 to October 1998, he was with the Optical Fiber Research Department, Bell Laboratories, Lucent Technologies, Murray Hill, NJ, working on a research project on multimode optical fibers for high-bit-rate local area network applications. His main research interest is in the investigation of advanced optical packet-routed systems, networks, and associated devices.

Mr. Düser is a student member of the IEE and the German Verean Deutscher Elektrotechniker (VDE). In 2000, he received the IEEE Lasers and Electro-Optics Society Graduate Student Fellowship.

Polina Bayvel (S'87–M'89–SM'00) received the B.Sc. and Ph.D. degrees in electronic and electrical engineering from University College London (UCL), London, U.K., in 1986 and 1990, respectively. During her Ph.D. work, she specialized in nonlinear fiber optics and its applications.

After a Visiting Research Fellowship at the Optical Fiber Department, Moscow General Physics Institute (USSR Academy of Sciences), she worked as a Principal Systems Engineer at STC Submarine Systems, Ltd., Greenwich, U.K., and Nortel Networks, Harlow, U.K., and Ottawa, ON, Canada, on the design and planning of high-speed optical fiber transmission networks, during 1990 to 1993. Currently, she is a Reader (Associate Professor) and a Royal Society University Research Fellow (1993–2003), at the Department of Electronic and Electrical Engineering, UCL, and heads the Optical Networks Group, UCL, whose research activities are in the areas of high-speed multiwavelength optical communication systems and novel wavelength-routed optical network architectures. She has authored and coauthored more than 120 journal and conference papers. She serves on the Editorial Board of the Optical Society of America (OSA) *Journal of Optical Networks* and on the technical program committee of the IEEE Lasers and Electro-Optics Society (LEOS) Annual Meeting. She is also an Honorary Editor of *Electronics Letters*. Since 2001, she has been the Royal Society representative on the European Science Foundation Committee for the Physical and Engineering Sciences (PESC).

Dr. Bayvel is a Fellow of the Institute of Electrical Engineers (IEE). In 2002, she was awarded the Paterson Medal and Prize from the Institute of Physics for her contributions to research in fundamental aspects of nonlinear optics and their applications in optical communications systems.