OPS or OBS in the Core Network?
- A Comparison of Optical Packet- and Optical Burst Switching

M. Nord (1,3), S. Bjørnstad (1,2), C. M. Gauger (4)
(1)- Telenor R&D, 1331 Fornebu, Norway; (2)-Norwegian University of Science and Technology (NTNU); (3)-Research Center COM, Denish Technical University; (4)- Institute of Communication Networks and Computer Engineering, University of Stuttgart, Germany.
martin.nord@telenor.com, steinar.bjornstad@fysel.stud.ntnu.no, gauger@ikr.uni-stuttgart.de

Abstract: Optical Packet Switching (OPS) and Optical Burst Switching (OBS) are optically switched networks architectures aiming to serve higher layer packet-based communication protocols. This paper gives a rationale for introducing these switching schemes in telecommunication networks, and discusses transparency and channel count in the context of contention resolution. Focusing on asynchronous operation with variable length optical packets/bursts, we review the latest work on OPS and OBS designs, contention resolution methods and OBS reservation mechanisms. The performance of both OPS and OBS is reviewed and held up against design complexity.

Key words: Optical packet switching, optical burst switching, contention resolution, packet-handling schemes, reservation mechanisms, burst assembly.

1. INTRODUCTION

1.1 Transparency of optical networks

OPS and OBS are optical network architectures that apply optical switching technology, and have an improved granularity compared to Optical Circuit Switching (OCS). OPS and OBS have different control plane solutions and typical packet/burst lengths. Electronic switching solutions require OEO conversions at every node, and thus knowledge of the signal bitrate and transmission format. Optical switching opens up for network transparency, i.e. design of a network that readily handles any format and bitrate. This is often cited as an attractive property of optical switching, together with the potential for high bit-rate systems (40 Gb/s or above). However, due to the required upgrade of old- and adding of new infrastructure, it is uncertain whether higher bitrates decrease the transmission system cost. Furthermore, realisation of true transparency is very hard: Increasing the bitrate or shifting from the NRZ to the RZ transmission format will increase the spectral occupancy of the signal. This increase may not be supported by neither demultiplexers, nor some switching fabrics such as array waveguide gratings (AWGs), because of their limited channel bandwidth. In addition, clock recovery techniques and optimum dispersion maps depend on the format and bitrate. Transparency will hence require complex active components in the network, whose cost is probably not justified by the gain in flexibility of input signal format and bitrate. Since it is the cost of the network as a whole that is important, one must include the cost of edge- and core nodes, transmission-, switching- and adaptation layers. The choice of contention resolution method is an important factor in the node cost, and suitable contention resolution is essential to have a network operating with high load and acceptable loss rates; available contention resolution methods are discussed in [1]. The importance of transparency should not be overrated when comparing OPS and electrical packet switching, as well as OPS and more hybrid solutions (e.g. OPS with electrical buffering as discussed later).

Many optical packet- and burst switch simulations, using partly wavelength domain contention resolution, reveal that high channel count dramatically reduces the need for buffering (e.g. [1]). We argue that high channel count with moderate bitrate channels is beneficial compared to a small number of high bitrate channels, when the wavelength domain as contention resolution method is preferred. However, this requires tuneable wavelength converters (TWCs). We assume TWCs to be available at a reasonable cost within a few years. Already, many packet switch designs assume widespread use of TWCs, both for contention resolution and for routing through an AWG.

1.2 Scope of study, OPS and OBS rationale,

Introduction of OPS or OBS in telecommunication networks should be motivated by improved performance, combined with reduced network capital and operational expenses. At present, with the immaturity of the technology and concepts, it is not possible to assess the cost accurately, but we will compare the complexity of OPS and OBS designs in later sections. A detailed comparison with electronic switching alternatives is also beyond the scope of this paper, however, we will here outline the motivation for moving from OCS to OBS and OPS: In OCS networks, all traffic of a connection follow the same path, so that the client packets arrive in the order they were sent. Transfers may suffer from long set-up times relative to connection holding time. Furthermore, since no buffering takes place, the capacity of the circuit must equal the peak data rate, which can be orders of magnitudes higher than the average data rate, for bursty sources. Hence, low utilisation results in networks with many communication pairs with bursty traffic patterns, as observed in the Internet.

To serve such a packet based network, OPS and OBS aim at improving the network utilisation by statistical multiplexing. OBS is furthermore suitable for transporting the entire data content of some types of communication in a single burst. Since set-up times are significantly reduced compared to OCS, a high adaptivity can be expected. The network could rapidly respond to fibre breaks and node break-downs.

Explicit transfer guarantees are not possible, since the limited transfer unit sizes prevent two-way reservation. Awaiting acknowledgement would lead to significant bandwidth waste, excessive delay and requirement for excessive buffering at the ingress node; even OBS has burst sizes inferior to return propagation delay in typical multi-hop core networks.

1.3 Packet/burst handling schemes

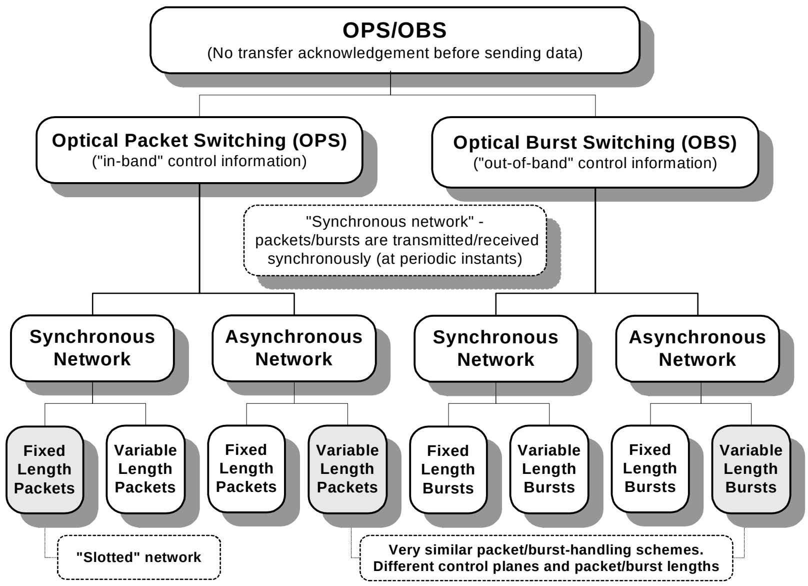

Figure 1. Potential and viable (grey boxes) OBS and OPS packet/burst handling schemes.

It is important to distinguish between the OPS and OBS network architectures and the packet/burst handling schemes. OPS and OBS networks can both operate with four different handling schemes, illustrated in Figure 1. These schemes were discussed in an OPS context in [1], and two schemes were identified as attractive, namely synchronous operation with fixed length packets (FLPs) (termed “slotted” scheme) and the asynchronous operation with variable length packets (VLPs).

A main motivation for OBS is to reduce the complexity of optical technology complexity. Therefore, avoidance of burst alignment is essential, thus the two synchronous schemes will not be studied here. To avoid padding and to give the burst assembly mechanism freedom to optimise burst lengths, which typically are based on timers, fixed burst length schemes are omitted.

The three most attractive OPS/OBS packet/burst handling schemes, resulting from this discussion, are marked with grey background in Figure 1. Each can make use of different contention resolution methods and OBS reservation mechanisms; hence the network design space is large.

1.4 Terminology

The term edge router (or ingress/egress routers) is commonly used. It can be unclear whether pure edge routers exist, or are merely invented for conceptual and illustrational purposes. We believe that most routers in the network core are edge routers in the sense that they will be situated close to a traffic aggregation point, and should thus have the possibility to aggregate packet/bursts from both higher layer clients and from neighbouring OPS/OBS domains. In this paper, edge router, corresponds to the OPS/OBS interface both to higher layers and to adjacent OBS/OPS networks.

1.5 Rationale for study

OBS and OPS receive a lot of attention in the research community. This paper provides an overview of recent achievements and analyses current trends in both fields. We observe that OPS using asynchronous packet handling scheme becomes more popular, and discuss proposals aiming to resolve the increased contention that results. Furthermore, we review OBS proposals including advanced reservation mechanisms, QoS differentiation methods, and recent work reflecting that buffering is required to maintain acceptable loss rates at reasonable loads. To conclude, we discuss and compare the complexity, performance, network scenarios and time-scales for OPS and OBS. We observe that OBS and OPS architectures tend to converge; there is therefore a need to analyse their similarities and fundamental differences.

No published study covers, to the best of our knowledge, this broad scope. Furthermore, earlier overviews assume slotted OPS, and typically do not include buffering for OBS, nor QoS differentiation for both schemes.

1.6 Outline

This introduction aimed to give the background, terminology and rationale for the study. In Section 2, we define the optical switching concepts, and compare their basic properties. In Section 3 and 4 we describe in greater detail advanced features of OPS and OBS, respectively. In Section 5, we compare OPS and OBS performance and complexity of each concept. We conclude our study in Section 6.

2. OPTICAL SWITCHING CONCEPT DEFINITIONS

OPS uses in-band control information; the header follows the rest of the packet closely, so there is no reservation possible. Reading and reinsertion of packet headers with strict timing requirements are required, due to the short packet duration- typically around 1 μs. Contention resolution is typically achieved by a combination of wavelength conversion, buffering and, in rare cases, deflection routing. Many different optical packet switch designs exist, a selection of latest achievements is reviewed in Section 3.

Due to the large amount of OBS approaches proposed to date, we define OBS's fundamental characteristics here, and leave the description and discussion of its different varieties to Section 4: OBS is a fast circuit switching technique providing granularity in between wavelengths and packets. Client layer packets are assembled in edge nodes, and transported through the optical network in optical bursts. A key characteristic is the hybrid approach: control information is signalled out-of-band using a control packet ("burst header") and processed electronically, while data bursts stay in the optical domain until they reach the egress node. Another key concept of OBS is one-pass reservation, i.e. burst transmission is initiated shortly after the burst was assembled and the control packet was sent out.

In Table 1 we compare basic properties of OCS, OPS and OBS to identify some differences. We assume that deflection routing is not used, and that forwarding is identical for packets/bursts, so that misordering may only result from buffering. Loss or connection set-up rejection may take place in all schemes. Latency is the end-to-end delay, which is minimised for OPS, where only negligible buffering adds to the propagation time. Buffering is typical for OPS, but is rarely used in conventional OBS schemes.

Table 1. Main properties of Optical Circuit-, Burst-, and Packet switched networks

	OCS (including set-up)	OBS (bufferless)	OPS (with buffers)
Suitable for data sizes of	> GB	~ tens of kB	~ 100-1500 B
Transfer Guarantee	Possible	No	No
Loss type	Set-up request rejection	Loss of burst	Loss of packet.
Control	Out-of-band	Out-of-band	In-band
Buffering	No	No	Typically
Latency given by	Set-up + propagation	Propagation + burst assembly	~ Propagation
Misordering	No	No	Possible
Control overhead	Connection set-up	Control burst channel	Packet header

3. OPS WITH VARIABLE LENGTH PACKETS

Recently, VLP OPS has gained much attention [1-4]. This approach is regarded as promising because of the continuing traffic growth in the Internet where the packet lengths are variable and not fixed. In this section, first buffering in OPS is discussed. Then, optical and electronic buffering of VLPs is discussed and compared. We then discuss how QoS differentiation based on differentiation on packet loss ratio (PLR), can be obtained both by electronic and optical buffering.

3.1 Minimising buffering in OPS

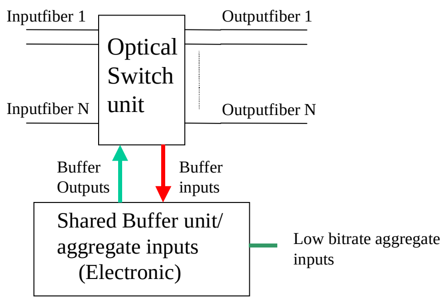

Figure 2. An optical packet switch architecture with shared buffering.

Buffering is one of the most demanding challenges in OPS. Optical random access memory is still immature, and optical buffering is therefore realised by using optical fibre delay lines, (FDLs). FDLs give a fixed delay, which depends on the fibre length. In a slotted scheme, it is found fare more efficient to use FDLs with different lengths, where all lengths are an integer multiple of the packet duration, than using fixed length FDLs [2]. The variable FDL length buffering scheme will however cause misordering of packets; there is no first-in, first-out (FIFO) packet scheduling.

When adding buffering to an optical packet switch, extra ports on the switching matrix will normally be needed. Both the added ports and the buffering itself increase the packet switch cost, so that the number of buffer interfaces should be minimised, as proposed in [1-3]. As illustrated in Figure 2, by sharing the buffer resources, the number of buffer interface inputs can be reduced, as in [1-4]. Because only a small amount of the packets passing through the switch will be buffered, the ratio of buffer input interfaces and input ports can be low, while still maintaining an acceptable PLR.

3.1.1 Buffering variable length packets using FDLs

Switching and buffering VLPs optically are however even more demanding than switching fixed length packets. Packets arrive asynchronously at the switch, causing less efficient utilisation of the switching matrix and a higher PLR. Using FDLs, a common approach is to assume that the FDL module introduces delays that are consecutive multiples of a delay unit 'D', like in [5]. In a buffer made with FDLs, the m'th FDL introduces a delay given by D_m= (m-1)D. The uniform distribution of delays is the simplest case, and can be considered as a sort of reference. Using FDLs for buffering VLPs is different from using an electronic buffer. In both cases, when a packet arrive at the switch input and cannot be served because of output contention, the packet will be buffered if vacant buffer resources can be found. When the output again becomes available, in the case of electronic buffers, the packet stored can immediately be served. However, when using FDLs, because of the discrete time steps, an additional delay caused by the waiting time before the packet arrive at the output of the FDL, may be introduced. The delay depends on when the packet was buffered, when it is being served, and the length of 'D', which also will influence the PLR. An optimal 'D' can be found that minimises the PLR, at a given load and mean packet length.

The performance of FDL buffering of VLPs is fairly poor [6]. A large number of FDLs is therefore necessary to achieve a low PLR. This can be implemented by using multiple stages of buffering, as analysed in [6]. In combiners at the output of each buffering stage of the analysed switch, the packets crossing the switch in a given instant are merged, they must therefore be coded on different wavelengths. For each incoming packet a suitable wavelength must be found. If no wavelength is available, the packet is dropped. Thus a limited pool of internal wavelengths is a set of resources where congestion may arise and cause packet loss. Also, because an approach with shared wavelength converters at the input is used, this will also be a source to packet loss. The third reason for a packet being dropped will be lack of buffering space. The analysis of a switch with a 16 X 16 switching matrix, shows that to achieve a low PLR, in the order of 10^{-6}, hundreds of different delays are necessary. This is not feasible with a single stage of FDLs, but can be implemented with consecutive stages of FDLs.

3.1.2 Electronic buffering

Since FDLs are known to be bulky, their use should be minimised, or eliminated, if possible. [1,3] propose to use an electronic buffer with a minimum number of buffer inputs. The buffer could be a FIFO buffer with a random access time of the next packet in the queue. Hence, buffered packets can be served immediately when an output becomes available, as opposed to the FDL case. A simple buffering policy is therefore viable. However, as discussed earlier, the transparency will be limited. Also, if very high channel bitrates (> 40 Gb/s) are used, the channel containing the packet to be buffered may need to be optically demultiplexed. OTDM techniques generate lower bitrate bitstreams that are divided on several electronic inputs. When using electronic buffers, the amount of memory, i.e. the buffer depth, will usually not be a limitation since electronic memory is inexpensive. However, each buffer interface requires an OE converter, which will add extra cost to the system. Therefore, it is important to minimise the number of buffer inputs, which also minimises the switch matrix port count.

3.2 QoS- using FDLs to buffer variable length packets

[7] exploits the wavelength domain to achieve service differentiation. This will be particularly useful for implementation of Assured Forwarding (AF), when using DiffServ. A packet switch design similar to the one analysed in [6] is studied. Three levels of service in terms of PLR are supported within a single AF class. These service levels are thought to correspond to high priority “in profile” traffic, high priority “out of profile” traffic and best effort traffic.

The aim is to obtain sufficiently low PLRs ($<10^{-6}$) for the highest priority traffic, and a difference of at least one order of magnitude for the remaining levels. Three techniques, that all exploit the characteristics of the switching architecture, are studied.

The first, wavelength allocation (WA), focuses on controlling contention on wavelength converters at the input. Service differentiation is obtained by differentiating the access to wavelength converters for each service level. The second principle combines wavelength allocation and threshold dropping (WA/TD). Wavelength converter access is differentiated in two classes, reserving some only for the highest level. Differentiation between the two lowest classes is achieved by means of a threshold dropping in the buffer. When the buffer occupation is above a threshold, the lowest priority packets are discarded. The third principle is called wavelength allocation with scheduling (WAS). Like in WA, different numbers of wavelength converters is assigned to each service level. In addition, if a high priority packet at the input of the buffer, does not find a wavelength to exactly obtain the required output queuing delay through the buffer, its delay is increased and the search repeated. The search can be repeated a maximum of three times, until it is discarded. For the second priority class only two repetitions are allowed, and for the lowest priority only one.

Performance analysis of the three techniques shows that, for a single buffer stage, the WA technique obtains a good differentiation of the PLRs, and that the benefits of using WAS are fairly small. This is shown to be different for a multistage architecture. The performance of the WA/TD technique, shows that priority management in the time domain is not as effective as in the wavelength domain, hence one of the before mentioned techniques are to be preferred. For a three-stage buffer, the performance of the WA technique is rather poor, but the WAS algorithm maintains a good performance. Hence, the analysis concludes that the WAS technique is the most suitable, and that imposing thresholds on the buffer occupancy is not necessary.

3.3 QoS by electronic buffers and variable length packets

In [3] an OPS transport network supporting two traffic classes is suggested. Service differentiation is suggested to be based on PLRs, since node delay is expected to be negligible in OPS networks. A PLR difference of three orders of magnitude is suggested, and the PLR for high priority traffic (HCT) should be 10^{-6}, which is expected to be sufficient for demanding real time services like MPEG-2. An evaluation of the suitability of supporting only two traffic classes in a transport network shows that only a small fraction of today’s network traffic will need to be HCT.

A packet switch design based on optical switching and a minimized number of electronic buffer input ports is suggested. A model with a node degree of 8 and a number of wavelengths per link higher than 32 is given as an example. Simulations show that the described optical packet switch design supports the required HCT PLR and service differentiation, by allocating a small fraction of the electronic buffer inputs to the HCT traffic. For both 32 and 128 link wavelengths, this is achieved by reserving only four out of a total of 32 buffer inputs. In a future network, a larger amount of demanding real time traffic requiring priority may be expected. The simulation shows that when the HCT traffic is 50 % of the total traffic, a larger number of reserved buffers are needed to obtain the same service differentiation..

3.4 FDL vs. electronic buffers for variable length packets

The papers studied show that switching asynchronously arriving VLPs is far more demanding than switching slotted packets. To compensate for the decreased switching efficiency, increased amount of buffering resources has to be applied. In addition, when FDLs are used for buffering VLPs, hundreds of different buffer delays should be available [7]. This can be implemented by using buffers with multiple stages. However, this increases the complexity compared to the slotted approach.

When using electronic buffers, theoretically an infinite number of delays are available. The highest delay available will depend on the maximum buffer depth, however this is not found to be a problem [1,3]. A low PLR is shown to be obtainable by employing a moderate number of OE buffer inputs, if the number of link wavelengths is 32 or higher.

Packet loss based service differentiation is obtainable both by using FDLs and by using electronic buffers. In the latter case, a simple reservation scheme, where a limited number of buffer inputs are reserved, is sufficient. When using FDLs, several reservation schemes were evaluated. A wavelength converter based reservation scheme, combined with assignment of wavelengths in the buffer, was found to be the most suitable. Simulations demonstrate successful QoS differentiation between three classes.

Under the assumptions made in the introduction, a high number of link wavelengths will be available in the future core network. Hence, by exploiting the wavelength domain for contention resolution, a low number of packets will need to be buffered.. We conclude that if full transparency is not important and if moderate channel bitrates that can be handled by electronics (year 2002: 40 Gb/s or less) are used, employing a minimised number of electronic buffer inputs may be attractive because of the high performance and simple implementation. If channel bitrate is too high to be handled by electronics, OTDM and subsequent demultiplexing can be employed. If transparency is important, FDLs will be attractive also for VLPs. Both buffering schemes demonstrate the capability of achieving what we think is a sufficiently low PLR, 10^{-6}, and the flexibility needed for supporting a class of service differentiation based on PLR.

4. OPTICAL BURST SWITCHING

OBS network performance depends on reservation scheme, burst assembly process, contention resolution method, QoS differentiation and node design. Here, we review recent OBS proposals on these topics.

4.1 Reservation mechanisms in OBS

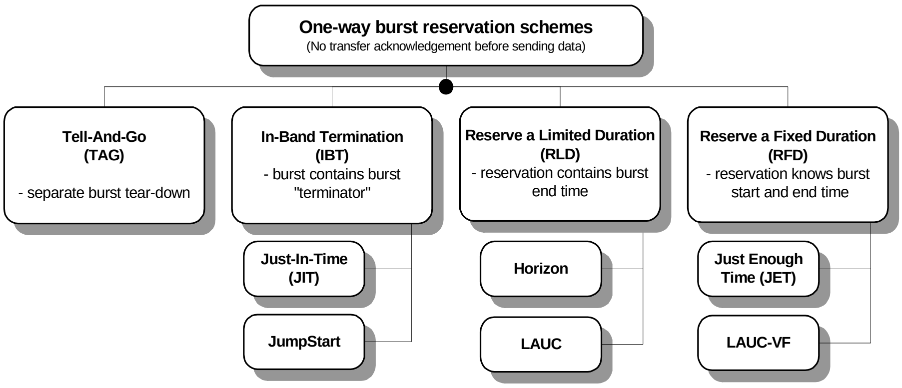

Figure 3. One-way reservation schemes, including prominent examples.

Time separation of control and data allows resource reservation for data burst transport such that the burst can cut through core nodes, and thus stay in the optical domain. Such reservation schemes have been the origin of OBS, and are rather well understood [8-10]. Figure 3 depicts a classification of one-way reservation schemes, and shows prominent examples.

4.1.1 Tell-And-Go and In-Band-Termination

In the Tell-And-Go (TAG) scheme, the source first sends a control packet on a separate control channel that immediately reserves bandwidth along a path for the following data. In case no resources are available the burst has to be blocked. From the edge, data is sent without awaiting acknowledgement. A “tear-down” control signal or packet, is sent to release the connection.

In-Band Termination (IBT) is very similar to TAG. The essential difference is that the end of burst transmission is not indicated by an explicit teardown packet, but by an in-band terminator. Thus, burst length is unknown until the terminator arrives. Detection of the IBT may be difficult, and loss of such a signal may lead to significant bandwidth waste. Due to their simple scheduling, the complexity of control units implementing TAG and IBT schemes are relatively low.

In general, the time between arrival of the control packet and of the data burst cannot be exploited by core switches to switch other bursts. Just-In-Time (JIT) [11] and JumpStart [12] are schemes that belong to this class.

4.1.2 Reserve a Limited Duration (RLD)

RLD mechanisms exploit the control packet's information on the burst length, and only reserve resources until the end of the burst. Thus, the offset time can overlap a preceding burst transmission as long as there is no reservation conflict among the data bursts. Horizon [13] and Latest Available Unscheduled Channel (LAUC) [14] have improved performance compared to TAG/IBT schemes. However, RLD mechanisms have higher complexity, and are harder to realise. Processing delay changes the delay between control packet and burst at each node. RLD and RFD (described below) thus require rewriting of the control packet offset field, unless fixed length FDLs are used to delay the bursts accordingly.

4.1.3 Reserve a Fixed Duration (RFD)

An even higher efficiency may be achieved if both the start and end times of accepted bursts are considered during reservation. A newly arriving burst can then be reserved in a gap left by already reserved bursts. RDF based schemes demonstrate reduced blocking in case of varying offset times [9], compared to other schemes. Just-Enough-Time (JET) [9, 15] and Latest Available Unused Channel with Void Filling (LAUC-VF) [14] are the most prominent RFD schemes. They have been extended for QoS support and further optimised, as discussed below. Recent proposals [16] employ larger offset times, and delay the reservation process until shortly before the transmission. They use this additional time for an optimised reordering of reservations. However, this can be expected to increase both complexity and delay. If fixed size FDLs are not used to prevent the processing related offset decrease at every node, JET suffers from an increased burst loss rate (BLR) of the bursts that approach the egress node (and thus have consumed more network resources). This can be inverted by increasing the offset time of long-travelling bursts, thereby increasing their priority [17].

4.1.4 Burst Segmentation/Composite Burst Switching

While all the latter schemes drop the entire optical burst if it cannot be reserved, schemes which are based on burst segmentation/composite burst switching only drop those parts of a burst that overlaps the already reserved burst [18-20]. Performance evaluation shows that this results in an improved (client) packet loss probability [18-21]. Drawbacks of these schemes are decreasing burst size in the network and increased effort for disassembling bursts at the egress node. Also, results in [18] show that these schemes require rather long initial burst sizes, which increases delay, and limited wavelength count to bring a significant improvement compared to pure wavelength contention resolution. Similar to the RFD schemes, QoS extensions have been proposed and will be discussed below.

4.2 Burst Assembly

Burst assembly is a key degree of freedom in OBS networks but also an additional functional block, which is not needed in OPS networks. As assembly defines burst traffic characteristics, it has significant impact on the performance of reservation mechanisms [22]. Assembly algorithms can be classified based on the decision criteria on when to send out optical bursts, and the order in which they are applied.

Time: Assembly schemes can decide on the transmission of a burst based on the maximum waiting time of the oldest packet in the assembly queue [14, 23] in order to keep delay requirements. This approach has been complemented by either padding to obtain a minimum size burst [24], or a token bucket traffic shaper in order to control the burst transmission process. In [22], upon an assembly time-out, the data volume exceeding the maximum burst size is sent out in a different burst, which is marked non-compliant and may therefore be dropped in the network. Time-based algorithms can better control delay requirements, but have to control unintended synchronisation, which might lead to periodic loss events [25].

Size: Assembly schemes which try to form bursts containing a certain data volume [26] or number of packets [27] can easily control burst length characteristics. They should be combined with a delay constraint in order to avoid excessive delays in low load situations. In [25], this problem is solved without a timer by adopting the assembly size based on offered load.

The influence of burst assembly on traffic characteristics [28] has motivated work aiming to exploit assembly in order to reduce the self-similar characteristic of IP traffic at the ingress of the optical network [24]. However, recent studies claim that burst assembly cannot yield lower self-similarity of data burst traffic (on a large time-scale) [29, 30].

4.3 QoS Supporting Schemes

In order to provide service differentiation directly in the optical layer several approaches have been proposed. Extending a classification in [22], we classify and discuss several approaches reported in literature:

Offset-based schemes: These schemes rely on the fact that a greater offset time translates into an earlier reservation, and thus in a higher probability of successful reservation. While the total blocking probability remains constant, high priority bursts can have significantly lower blocking. Prioritized JET (pJET) [31] is based on this principle, but has the drawback of sensitivity with respect to burst characteristics of low priority bursts and interference with basic offset adaptation [32]. Also, in combination with FDL buffers, the QoS offset has to be carefully adapted to the FDL delays to not deteriorate performance. In order to control the effects of these drawbacks, either additional strategies have to be designed or additional intelligence, e.g. global network view, has to be provided.

Preemption-based schemes: In these schemes, high priority bursts can preempt bursts of lower priority in case of an reservation conflict [33]. In case of burst segmentation, burst priority decides which of the conflicting bursts will be segmented [34]. If a high priority burst conflicts with an already accepted low priority burst, this segmentation can be interpreted as partial preemption. When offsets are applied, preemption may have the drawback that pre-empted burst resources are either wasted in downstream nodes, or that these nodes have to be informed of the pre-emption. This further increases signalling load in high load situations.

Dropping-based schemes: Similar to schedulers in the electronic domain, which provide proportional service differentiation by intentionally dropping packets in order to maintain a certain loss probability, some approaches provide service differentiation by actively dropping bursts. In [35], a certain ratio of loss probabilities is strictly maintained on short time-intervals which leads to a high total blocking probability. In contrast, [36] marks bursts as compliant/non-compliant according to their compliance with a pre-specified traffic envelope. During high load situations, non-compliant bursts are dropped. By over-provisioning, high priority classes are assigned a greater envelope, which yields fewer non-compliant bursts, and thus a lower loss probability. Dropping-based schemes are rather simple to realize although they require the additional functional block of a dropping unit.

Rescheduling-based schemes: As mentioned, it has been proposed to increase the number of burst control packets available for rescheduling, by a larger basic offset time and by further delaying the reservation decision in core nodes [36, 37]. These schemes require a more complex reservation control unit. Also, it can be expected that these schemes will lead to increased delay and lose effectiveness in larger networks.

Access-restricted schemes: These do not perform complete sharing but restricts access to resources based on class priorities. In [38], a scheme is proposed which maintains a database of recently used wavelengths and probes only a number of idle wavelengths depending on the QoS class.

4.4 Contention Resolution in OBS

Although advanced reservation schemes may be implemented, blocking still occurs due to statistical multiplexing. An OBS node that is not able to handle an incoming burst reservation will delete the burst when it arrives– hence retransmission may be initiated by a higher layer protocol, such as TCP. Retransmission in the optical layer, would probably be too complex, considering the enormous amount of data that would need to be stored.

Contention resolution can use the time, wavelength and space domain to minimise loss, just like in OPS, and the chosen method influences delay, burst misordering and acceptable network load for a given BLR. When using buffering for contention resolution, one should consider using electronic buffers (as proposed in OPS). It is easier to exploit later vacant periods in reservations when buffered bursts are randomly accessible. Contention resolution with respect to OBS specialities, e.g. an offset-based QoS reservation scheme, has been studied e.g. in [39-42]. Contention resolution in combination with reservation mechanism is needed when the BLR requirements of the different service classes are rather strict and when operating at a relatively high load.

4.5 Node architectures in OBS

As OBS and asynchronous OPS have very similar data plane requirements several results reported in literature on the design of optical nodes can be applied to OBS. Recently, optimised OBS node architectures have been designed [43, 44] and their scalability has been studied [45]. Both SOAs [44, 45] and AWG based architectures [43] are considered for OBS.

5. COMPARISON OF OPS AND OBS

5.1 OPS vs. OBS logical performance

In this paper we compare only fundamental properties like the delay and complexity. This is because it is very hard to compare logical performance of published work, since there always will be variations in the number of links and wavelengths per link, node architectures, scheduling and OBS reservation schemes. Furthermore, the offered load and the arrival characteristics of packets often varies; these parameters have an significant impact of the performance of any packet/burst switch. Different burst assembly mechanisms also modifies the input traffic characteristics. Finally, there are differences in how the loss/throughput is measured, as studies consider either client packets or optical packets/bursts. In the latter cases the overhead must be taken into account for relevant comparisons. Optical guard bands between packets occur more frequently than between bursts, due to the limited packet length. However, OBS may still suffer from a significant overhead, if the bandwidth of the OBS control channel is not negligible to that of the channels carrying the associated bursts.

We have however observed that recent simulations of packet/burst switches show low packet loss rates, when employing the wavelength domain and buffering in both schemes. Combined with the demonstrated QoS differentiation, this indicates that both OPS and OBS have the potential to meet future network requirements, but further work is needed to quantify the benefits of the switching schemes. At present, there is a lack of *network* simulations applying a node degree and a number of WDM links representative of a realistic future optical core transport network. Such a study would require massive modelling and simulation work, but may be required for more accurate assessment of networks.

5.2 OBS vs. OPS delay

We assume that no packets or bursts are lost, so that retransmission delay is neglected. The end-to-end delay experienced by client packets, measured from the moment it arrives at the OPS/OBS edge router, is therefore the sum of propagation delay, edge delay and the core node processing delay. The former is the same for OBS and OPS, and corresponds to approximately 0.4 ms per 80 km link. Propagation delay will therefore be the largest factor, but it is constant. The jitter, or delay variation, is another important factor and has origin in the edge and processing delay.

Processing delay is added at each hop. It includes the delay that each packet/burst experiences in the FDL for processing purposes, the time spent in switch matrix and in the buffers. Misordering of packets/burst may result, but can be avoided if buffer delay is limited or by implementing FIFO scheduling mechanisms.

Edge-delay is significant in OBS. The maximum time spent in the edge buffer is typically the sum of the time the first packet in the burst must wait until the last one arrives and the time offset between control packet and burst. Therefore, it depends on the burst size, the arrival rate of client traffic bound for that destination and the reservation mechanisms. A per-queue timer puts a boundary to the edge delay. The edge delay may be reduced if burst transmission starts before the burst is completely aggregated, as discussed in the context of pJET in [46]. OPS typically has much lower edge delay, due to its absence of offset, and smaller packet size, which limits assembly time.

5.3 OPS vs. OBS complexity

It seems more complex to encode, read and reinsert OPS headers than burst headers: Timing requirements are much more strict in the former case, and the reinsertion of headers must follow the packet's new output port and wavelength. However, the OPS community is investigating alternatives to serial headers such as Sub-Carrier Multiplexing (SCM) to ease the timing requirements and reduce overhead. OBS typically uses only a small number of wavelengths for control, compared to the data channels, but it is a challenge to maintain correlation between control packets and bursts.

Due to variations in the minimum durations of packets and bursts, the demands on the switching time of the switching matrix will vary. Required switching time depends on the acceptable overhead, transmission rate, packet/burst format and header encoding, but a rough estimate can be obtained by considering typical payloads: IP packets with 500 bytes have payload durations of 400 ns at bit rates of 10 Gb/s. If a *switching* overhead of 5 % is accepted, this corresponds to 20 ns switching time. At the same threshold value, a burst size of 10 kB gives required switching times of 400 ns. Hence, OBS can accept a slower switching matrix in the core nodes.

Common for OPS and OBS is that wavelength conversion for contention resolution typically is necessary, and that buffering in the core nodes, either by FDLs or electronic buffers will, except in rare cases, be required.

Edge router design differs between OPS and OBS when it comes to size and sophistication of electronic buffering systems. Each OBS ingress node needs at least one queue per egress node, each implemented in a separate electronic buffer to allow for random and parallel wavelength allocation. The output clocking speed of the buffer must equal the core bitrate, which is challenging above 10 Gb/s. In addition comes the control unit that implements the burst assembly and queue fairness algorithms. OPS sometimes apply a small amount of aggregation to limit the overhead associated with packet format, but the associated electronic buffers will be limited in size compared to OBS. Furthermore, direct encapsulation of single client layer packet is an option, eliminating the need for such queuing systems. This opens up for OPS closer to the users (MANs and access networks). OBS, on the other hand, requires larger amount of traffic for the burst assembly, to avoid either excessive latency or small burst sizes.

6. CONCLUSION

We have given an overview of the latest achievements in OPS and OBS, and studied how their differences affect the network performance and complexity. The discussion on network context challenges the research community's traditional requirement of truly transparent networks. Also, we stress the importance of high wavelength count systems to potentially reduce overall cost in networks with statistical multiplexing, which counters the current trend of high bitrate systems. If true network transparency is found too expensive, one should consider using electronic buffers. Results from work on optical packet switches employing electronic buffering reveal good performance with a relatively low number of buffer ports.

Traditionally, OBS assumes no buffering in the core. However, when operating at reasonable loads, it is observed that buffering in core nodes is required to match the maximum loss requirements of strict real time services, like MPEG-2. Due to the aggregation and reservation mechanisms available for OBS, the need for buffering is still limited compared to OPS. OPS traditionally have used a slotted scheme. However, motivated by avoidance of packet alignment, packet fragmentation and reduction of overhead, recent work also focuses on asynchronous, VLPs. This has increased the need for buffering, and due to the asynchronous operation, alternatives to FDLs have been considered.

There seems to be a convergence of the OPS and OBS concepts, when it comes to packet/burst handling and contention resolution. Acceptable loss rates can be achieved in both concepts, but further work is required to quantify these factors; the importance of similar simulation parameters in such a study was stressed. However, it is clear that both OPS and OBS can support QoS differentiation. Some differences remain: The OBS burst assembly reduces the need for optical buffering and may smoothen out self similar traffic, although there are different views on this subject. In addition, the reservation mechanisms give OBS an increased design space, and OBS can be realised with slower optical switching technology. This comes at the expense of increased edge node complexity, increased latency and jitter. The intermediate granularity limits OBS scenarios to core networks. OPS has wider potential, but requires that the technical challenges of fast switching and header reading and reinsertion can be solved at a reasonable cost.

In general, we see OBS as a first step towards a real dynamic network, with OPS as the ultimate goal. Both OPS and OBS receive a lot of attention, and we expect that the technology to make large-scale prototypes will evolve and become available within 3-10 years.

ACKNOWLEDGEMENTS

M. Nord and S.Bjørnstad would like to thank Telenor R&D and The Research Council of Norway for Ph.D. funding, as well as their supervisors, L. Dittmann at COM, and D.R. Hjelme at NTNU, respectively.

REFERENCES

1. S. Bjørnstad, M. Nord, D.R. Hjelme, N. Stol. Design of optical packet switched networks. Submitted *to Working Conference on Optical Network Design and Modelling* (ONDM 2003), Budapest.
2. C. Develder, M. Pickavet, P. Demeester. Choosing an appropriate buffer strategy for an optical packet switch with a feed-back FDL buffer. In: *Proc. of European Conference on Optical Communication 2002 (ECOC 2002).* Copenhagen, Denmark.
3. S. Bjørnstad, N. Stol, D. R. Hjelme. Quality of service in optical packet switched systems. In: *Proc. of SPIE Asia Pacific conference on Optical Communication 2003 (APOC 2003)*. Shanghai, China.
4. L. Tancevski, A. Ge, G. Castanon: Optical Packet switch with partially shared buffers: Design principles. In: *Proc. of Optical Fibre Communication conference (OFC 2001).* TuK 3, Anaheim.
5. F. Callegati. On the design of optical buffers for variable length packet traffic. In: *Proc. of Computer Communications and Networks 2000*. Pp. 448 –452.
6. F. Callegati, G. Corazza, C. Raffaelli. Design of a WDM optical packet switch for IP traffic. In: Proc. of *IEEE Globecom 2000. Pp.1283-1287, San Fransisco, USA.*
7. F. Callegati, G. Corazza, C. Raffaelli. Exploitation of DWDM for optical packet switching with quality of service guarantees. In: *IEEE Journal on Selected Areas in Communications (JSAC).* 20(1), 190-201, January 2002.
8. C. Qiao, M. Yoo. Choices, features and issues in optical burst switching. In: Optical Networks Magazine. 1(2), 37-44, April 2000.
9. K. Dolzer, C.M. Gauger, J. Späth, S. Bodamer. Evaluation of reservation mechanisms for optical burst switching. In: *AEÜ International Journal of Electronics and Communications.* 55(1), Jan 2001.
10. A. Detti, M. Listanti. Application of tell & go and tell & wait reservation strategies in a optical burst switching network. In: *Proc. IEEE International Conference on Telecommunications 2001 (ICT 2001).* Volume I-IV, GEOMA, Bucharest.
11. A. Detti, V. Eramo, M. Listanti. Performance Evaluation of a New Technique for IP Support in a WDM Optical Network: Optical Composite Burst Switching (OCBS). In: *IEEE Journal of Lightwave Technology (JLT).* 20 (2), 154-165, February 2002.
12. J.Y. Wei et al. Just-in-time optical burst switching for multiwavelength networks. In: *Proc. of the International Conference on Broadband Communications 1999.* Pp. 339-352, Hong Kong.
13. I. Baldine, G. Rouskas, H. Perros, D. Stevenson. JumpStart: A just-in-time signalling architecture for WDM burst switched networks. In: *IEEE Communications Magazine.* 40(2), 82-89, February 2002.
14. J.S. Turner. Terabit burst switching. In: *Journal of High Speed Networks.* 8(1), 3-16, January 1999.
15. Y. Xiong, M. Vanderhoute, C.C. Cankaya. Control architecture in optical burst-switched WDM networks. In: *IEEE JSAC.* 18(10), 1838-1851, October 2000.
16. C. Qiao. Labeled optical burst switching for IP-over-WDM integration. In: *IEEE Communications Magazine.* 38(9), 104-114, September 2000.
17. J.-B. Chang, C.-S. Park. Efficient channel-scheduling algorithm in optical burst switching architecture. In: *Proc. of Workshop on High Performane Switching and Routing 2002 (HPSR 2002).* Session 4B-2, Kobe, Japan.
18. B.-C. Kim. Performance of optical burst switching techniques in multi-hop networks. In: *Proc. of Globecom 2002.* Taipei, 2002.
19. V.M. Vokkarane, J.P. Jue. Burst segmentation: an approach for reducing packet loss in optical burst switched networks. In: *Proc. of the IEEE International Conference on Communications (ICC 2002).* Pp. 2673-2677, New York City, USA.
20. A. Maach, G. von Bochman. Segmented Burst Switching: Enhancement of Optical Burst Switching to decrease loss rate and support quality of service. In: *Proc. of ONDM 2002.* Torino, Italy.
21. M. Neuts, H.V. Vu, M. Zukerman. Performance analysis of optical composite burst switching. In: *IEEE Communications Letters.* 6(8), August 2002.

22. K. Dolzer. Assured Horizon – A new combined framework for burst assembly and reservation in optical burst switch. In: *Proc. of the European Conference on Networks and Optical Communications 2002 (NOC 2002).* Darmstadt, Germany.
23. L. Tancevski, S. Yegnanarayanan, G. Castanon, L. Tamil, F. Masetti, T. McDermott. Optical routing of asynchronous, variable length packets. In: *IEEE JSAC.* 18(10), 2084-2093, October 2000.
24. A. Ge, F. Callegati, L.S. Tamil. On optical burst switching and self- similar traffic. In: *IEEE Communications Letters.* 4(3), 98-100, March 2000.
25. S.Oh, H. Hong, M. Kang. A data burst assembly algorithm in optical burst swithcing networks. In: *Electronics and Telecommunication Research Institute (ETRI) Journal.* 24(4), 311-322, August 2002.
26. L.B. Softman, T.S. El-Bawab, K. Laevens. Segmentation overhead in optical burst switching. In: *Proc. of the SPIE OptiComm 2002.* Pp. 101-108, Boston, USA.
27. V.M. Vokkarane, K. Haridoss, J.P. Jue. Threshold-based burst assembly policies for QoS support in optical burst-switched networks. In: *Proc. of the SPIE OptiComm 2002.* Pp. 125-136, Boston, USA.
28. K. Laevens. Traffic characteristics inside optical burst switched networks. In: *Proc. of the SPIE OptiComm 2002.* Pp. 137-148, Boston, USA.
29. X. Yu, Y. Chen, C. Qiao. A study of traffic statistics of assembled burst traffic in optical burst switched networks. In: *Proc. of the SPIE OptiComm 2002.* Pp. 149-159, Boston, USA.
30. G. Hu, K. Dolzer, C.M. Gauger. Does burst assembly really reduce the self-similarity? In: *Proc. of OFC 2003.* Atlanta, USA.
31. M. Yoo, C. Qiao, S. Dixit. Optical burst switching for service differentiation in the next-generation optical internet. In: *IEEE Communications Magazine.* 39(2), 98-104, February 2001.
32. K. Dolzer, C.M. Gauger. On burst assembly in optical burst switching networks- a performance evaluation of Just-Enough-Time. In: *Proc. of the 17th International Teletraffic Congress (ITC 17).* Pp. 149-160, Salvador da Bahia, Brazil, September 2001.
33. D. Liu, M. Liu. Priority-based burst scheduling scheme and modeling in optical burst-switched WDM networks. In: *Proc. of ICT 18.* June 2002.
34. V.M. Vokkarane, J.P. Jue. Prioritized routing and burst segmentation for QoS in optical burst switched networks. In: *Proc. of OFC 2002.* Pp. 221-222, Anaheim, USA.
35. Y. Chen, M. Hamdi, D.H.K. Tsang. Proportional QoS over OBS Networks. In: *Proc. of IEEE Globecom 2001.* San Antonio, USA.
36. M. Yang, S.Q. Zheng, D. Verchere. A QoS supporting scheduling algorithm for optical burst switching DWDM networks. In: *Proc. of Globecom 2001.* San Antonio, USA.
37. D. Liu, M. Liu. Differentiated services and scheduling scheme in optical burst-switched WDM networks. In: *Proc. of the International Conference on Networks (ICON 2002).*
38. T. Hashiguchi et al. CoS-oriented wavelength assignment algorithm in burst switching optical networks. In*: Proc. of Opto-Electro and Communication Conference 2001*. (OECC 2001). Sydney.
39. M. Yoo, C. Qiao, S. Dixit. QoS performance in IP over WDM networks. In: *IEEE JSAC.* 18(10), 2062-2071, October 2000.
40. C.M. Gauger. Performance of converter pools for contention resolution in optical burst switching. In: *Proc. of the OptiComm 2002.* Boston, USA.
41. C.M. Gauger. Dimensioning of FDL buffers for optical burst switching nodes. In: *Proc. of ONDM 2002*. Torino, Italy.
42. R.-B. Yang et al. Dynamic thresholds for shared buffer switches with variable length packets. In: *Proc. of HPSR.* Session 3B-1, Kobe, Japan.
43. J. Ramamirtham, J. Turner. Design of wavelength converting switches for optical burst switching. In: *Proc. of Infocom 2002. Vol 1, pp. 362-370,* New York City, USA.
44. H. Feng, E. Patzak, J. Saniter, Size and cascadability limits of SOA based burst switching nodes. In: *Proc. of ECOC 2002.* Copenhagen, Denmark.
45. H. Buchta, C.M. Gauger, E. Patzak, J. Saniter. Limits of effective throughput of optical burst switches based on semiconductor optical amplifiers. In: *Proc. of OFC 2003*. Atlanta, USA.
46. C. Qiao, M. Yoo. Optical burst switching (OBS) - a new paradigm for an optical Internet. In: *Journal of High Speed Networks.* 8(1), 69-84, January 1999.