The cost of transmission in fibre-optical networks has been much reduced in recent years. With increasing bandwidth demand, the cost and speed limitations of electronic switching are expected to become the limiting factors in future telecommunication core networks. Optical packet switching (OPS) and optical burst switching (OBS) have the potential to remove this bottleneck, by introducing statistical multiplexing in optical switched networks. Optical hybrid switching combines optical packet or burst switching with Optical Circuit Switching (OCS), and, thus offering both high throughput and a circuit switched quality. In this paper we give the rationale for introducing OPS, OBS or a hybrid network and describe the required switching functionality in each case. Then we discuss and analyse different packet/burst handling schemes and make a detailed simulation analysis of an optical packet switch, identifying the most promising design choices. Furthermore we describe a hybrid switch and present simulations showing its basic properties.

Introduction

Improved residential access and corporate communication needs may trigger an increased demand for core networks supporting multimedia applications such as video-conferences, online gaming, video-on-demand and web-television streaming. Supporting these services requires a network with a better performance than today’s IP network. The future “layer 3” network will benefit from being served by a high-capacity optical layer network that complies with the following requirements:

1. Support high utilisation of resources; i.e. handle high capacity and link loads in a cost efficient way;
2. Support fine granularity that can be tailored to the applications’ needs;
3. Support the quality needed by strictly real time services; i.e. low packet loss and low delay;
4. Support variable length IP packets, as used in the Internet today.

We will start by giving a background for the current developments in technologies for core networks, before discussing how the different switching techniques have the potential to meet these requirements.

In the mid-1990s, the introduction of the erbium doped fibre amplifier (EDFA) enabled the more bandwidth efficient wavelength division multiplexing (WDM) transmission systems that resulted in reduced core transmission cost. However, cost-effective core networks also require cost-effective switching and routing. Today, the switching is performed electronically; optical to electrical to optical (O/E/O) conversions are hence required in the switching nodes. With increasing bandwidth demand, the cost of these conversions and the limited capacity of electronic switching may become a network capacity bottleneck. Implementing optical switches with ms-range switching time, e.g. based on micro electro-mechanical systems (MEMS) technology, is nowadays becoming feasible. Currently, operators and vendors are working on control plane architectures for resulting optical circuit switched (OCS) networks. OCS networks offer explicit transfer guarantees. This leads however to set-up delays of at least several milliseconds, due to the propagation delay induced by awaiting the set-up confirmation.

Future optical networks should be able to serve a client layer that includes packet-based networks, such as the Internet. These may have a highly dynamic connection pattern, with a significant portion of bursty traffic between the communicating pairs. In this case, OCS transport at a granularity of 2.5 Gbit/s, 10 Gbit/s or 40 Gbit/s may lead to low resource utilisation. It would require over-dimensioning of the number of connections as well as of the bandwidth reservation of each connection, in order to avoid excessive delay and extensive buffering at the ingress router. Optical packet switching (OPS) and optical burst switching (OBS) have the potential to overcome these problems by introducing statistical multiplexing (SM) at the optical layer. Several optical switching technologies have demonstrated nanosecond to microsecond range switching times, opening up for these new switching schemes. Telenor R&D and NTNU participate actively in European based studies of both concepts, through the ongoing IST projects NOBEL [1] and STOLAS [2], the recently concluded COST 266 action “Advanced Infrastructure for Photonic Networks” [3], as well as the ongoing COST 291 action “Towards digital optical networks”. Fur-
thermore, a hybrid concept combining packet switching, optical or electronic, with optical circuit switching, is studied in the Optical Packet switched Migration capable Network with Service Guarantees (OpMiGua) project. The circuit switched part of the network offers a Guaranteed Service Transport (GST), while the Statistically Multiplexed (SM) part of the network ensures high throughput for a Best Effort (BE) type of traffic. The OpMiGua project receives parts of its funding from the Norwegian national research council (NFR). Project partners are Telenor, NTNU and Network Electronics, a Norwegian company developing and manufacturing WDM systems for broadcasting applications.

OPS and OBS concepts
Inevitably, there are some differences in terminology within the research community. We start this section by explicitly describing some concepts and terms used in this article. Both OPS and OBS are based on the idea of separating forwarding from switching in the network nodes. Forwarding decisions are based on the information contained in the optically transmitted control information. As illustrated in Figure 1, in the OPS case this is realised as a packet header while in the OBS case it can be a separate packet preceding the burst (Burst Control Packet, BCP). The BCP may travel together with the burst at a separate wavelength, alternatively it is encoded on the burst using a different modulation format than used for the payload, e.g. as in IST STOLAS [2]. The BCP and OPS header are converted to the electronic domain at the packet/burst switch input interface. It is subsequently electronically processed in the control unit, which then configures the node resources for switching, wavelength conversion or buffering of the data unit. On the other hand, the burst or packet payload is optically switched, thus avoiding the costly O/E/O conversion, and in principle simplifying the interface cards to the optical fibres.

OPS definition
OPS assume in-band encoding of control information (same wavelength, simultaneous transmission). Since the header follows the rest of the packet closely, 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. Typically, OPS proposals assume packet sizes from around 40 to 1500 B, representative of today’s Internet traffic [4]. Some aggregation of client packets to form an optical packet may be required, since the smallest client packet sizes only correspond to 32 ns payload at 10 Gbit/s bitrates. The relation between packet format, switching technology and overhead is studied in more detail in [5].

OBS definition
OBS was introduced quite recently [6] and aims at offering granularity in between optical wavelength channels and optical packets, at the same time minimizing optical complexity. OBS assumes more extensive burst aggregation to form bursts with tens of kB of payload, thereby relaxing the requirement to optical switches, from the nanosecond to the microsecond range. Client layer packets are assembled in edge nodes, and transported through the optical network in optical bursts. Another key concept of OBS is one-pass reservation; i.e. burst transmission is initiated without awaiting a set-up confirmation. As illustrated in Figure 1, BCPs are transmitted ahead of the corresponding burst. The relative delay is termed offset, and as a minimum it equals the time required to settle the switches and perform control processing in the network nodes. The reservation scheme depends on the amount of information that the BCP contains on its corresponding burst. Reserve a Fixed Duration (RFD) type BCPs include both the start- and end times of bursts. The combination of RFD scheme, offset and the principle of delayed reservation (DR), enables advanced burst scheduling. For example, a newly arriving burst can be reserved in a gap left by already reserved bursts. This can optimise bandwidth usage and enable offset based QoS differentiation methods. A more detailed overview of OBS is provided in [7].

Optical Circuit and Packet switched Hybrid (OpMiGua)
When SM is efficiently exploited, constant delay and zero packet loss is not guaranteed. Combining properties of Optical Circuit Switches (OCS) with those of optical packet switches (OPS), may offer both better cost and performance [8]. Recently, a hybrid network concept was proposed [9, 10], combining packet switches with a Wavelength Routed Optical Network (WRON); i.e. an optically circuit switched network as shown in Figure 2. In this concept, a frac-

![Figure 1 Main differences for transferring control information in OBS and OPS. a) illustrates that OBS typically has an “out-of-band” BCP, transmitted on a different wavelength and with a time-offset compared to the burst payload. b) illustrates that OPS typically has an “in-band” header, transmitted at the same wavelength and simultaneously as the packet payload. A number of variations on these main principles have been proposed](image-url)
tion of the packets, the Guaranteed Service Transport (GST) packets, follow the wavelength paths defined by the configuration of the optical cross connects (OXC)s. The remaining packets (SM packets) are switched in packet switches according to their header information. The GST packets are not subject to packet loss, and delay is fixed. On the other hand, SM packets have a BE quality. They will be dropped during congestion, and the use of buffering implicates a variable delay.

In the hybrid network concept the link resources are divided in the time domain with packet length granularity. By sending GST packets through the WRON, and sending lower quality packets through the packet switches on link bandwidth not consumed by GST traffic, both high throughput efficiency and the guaranteed service are achieved [8]. The GST class is able to accommodate services requiring circuit switched quality, while lower quality traffic is statistically multiplexed and experience only a moderate penalty from the GST traffic. We term this network concept Optical packet switched Migration capable network with Guaranteed service (OpMiGu). If the WRON network is used for transit traffic, studies show that hybrid networks achieve potentially large savings compared to both packet- and circuit switched networks [11]: Compared to a circuit switched network, one can save wavelengths because SM improves the utilization of the wavelength paths. Compared to a packet switched network, reduction in packet switch sizes can be achieved, because the packet switches do not process packets that are sent through the WRON. Assuming a reference network applied within the COST 266 project with associated traffic demand forecast for 2008, it is found that the hybrid network only needs 1/3 of the wavelengths needed in a pure circuit switched network. Compared to a pure packet switched network, the amount of packet processing in the nodes is reduced corresponding to a 30 % reduction in number of wavelengths.

**Comparison of optical switching schemes**

In Table 1 we compare basic properties of OCS, OPS and OBS, aiming at identifying main properties of each switching paradigm. Note that these concepts are not standardised, and no universally accepted definition exists.

Recent papers on OPS/OBS show a converging trend: First, the packet/burst handling schemes discussed below are becoming more similar. Second, in contrast to conventional OBS, buffers are now being implemented in new OBS designs, similar to OPS designs. Note that implementing buffers imply that packets may arrive out-of-order even when packets in the same stream follow the same network path. Such out-of-order delivery is avoided in OCS. However, some important differences remain. These include finer granularity, but also higher overhead in OPS. OBS still assumes out-of-band control encoding and more advanced scheduling. Furthermore, its coarser granularity gives more tolerance to the switching time of switch matrices.

In OCS, since each connection has reserved dedicated resources, it can offer explicit transfer guarantees. This means that once a connection is set up, there will be no loss of packets inserted on that circuit. In contrast, such explicit/absolute transfer guarantees require retransmission protocols for OPS and OBS networks. However, properly dimensioned networks, possibly combined with “MPLS type” circuit oriented traffic engineering, can offer statistical quality of service guarantees which may serve the purpose, e.g. average packet loss rate (PLR) of 10⁻⁶. Moreover, employing QoS differentiation gives the network designer the possibility of offering two or more PLR levels to the clients.

![Figure 2 Hybrid network model illustrating the efficient sharing of the physical layer. If the WRON network is an S-WRON, the cross connect can be a matrix, manually configurable. If a D-WRON network is preferred, the cross connect should be configured by a control plane. The connected cross-connects and packet switches are physically co-located. These can be separate units with a common control unit, or they can be integrated, sharing physical resources in the node.](image-url)
The hybrid scheme combines absolute and statistical QoS guarantees. Headers are only needed on the BE type of traffic. If a Static WRON (S-WRON) is employed, overhead for setting up connections is avoided for the GST packets. GST connections may be of the duration of a packet, burst or dynamic circuit. In this paper we consider the connection duration to correspond to that of a burst.

**Packet/burst handling schemes**
Both in OBS and in OPS the switching matrix is required to be re-configurable on the burst/packet time scale, to allow the packets/bursts to efficiently share node and fibre resources. The switching operation is hence more demanding in OPS than in OBS. The basic principles for switching architectures and functionalities are in principle independent of the packet/burst-handling scheme. In the hybrid scheme, the BE traffic may be burst switched or packet switched, thereby having the same requirements to the switching architectures as OBS and OPS, respectively. The GST packets follow circuit switched paths, hence demands to the switching matrix for these packets are equal to those for circuit switching. In Figure 3 we show the four potential handling schemes for OPS and OBS, classified according to synchronisation and data unit type.

The following summarises main points to consider when making a choice concerning operation mode and data unit type (fixed or variable length).

In asynchronous optical networks, each transmitter/output port in the nodes emits optical packets asynchronously, so that packets arrive at random moments at the OPS nodes’ interfaces. Synchronous operation requires optical synchronisers for packet alignment at the switch interfaces and a global network clock for practical realisation.

Fixed length packets (FLP) schemes require fragmentation of client packets and padding to fill the optical packets. Both factors increase the overhead; i.e. bandwidth consumption for successful transmission through the OPS network [5], and thus the blocking ratio for the same client load. In addition, fragmentation calls for reassembly of client packets, which increases egress node complexity. For fragmented packets, it can in some cases be hard to tell during reassembly whether a missing fragment is lost or simply delayed. Fibre delay line (FDL) based buffer design and management is simpler for FLP than for variable length packets (VLP), and it is simpler to maintain the packet order. The complexity of other contention resolution methods may be independent of the packet length. In general, a switch matrix operating in synchronous, FLP mode will have less contention than when operating in asynchronous mode [12]. Furthermore, in this mode, a re-arrangeable non-blocking switch matrix may have equal performance to a strictly non-blocking switch matrix [13]. For optimum scheduling, variable packet length requires coding of packet duration in the packet header.

---

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

<table>
<thead>
<tr>
<th></th>
<th>OCS (including set-up)</th>
<th>OBS</th>
<th>OPS</th>
<th>Hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommended size of transfer unit</td>
<td>&gt; GB</td>
<td>- tens of kB</td>
<td>- 40 – 1500 B</td>
<td>- 40 B – &gt; GB</td>
</tr>
<tr>
<td>Explicit transfer guarantee</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes (GST part)</td>
</tr>
<tr>
<td>Statistical QoS guarantees possible</td>
<td>N/A</td>
<td>Yes</td>
<td>Yes</td>
<td>N/A (GST part) Yes (BE part)</td>
</tr>
<tr>
<td>Loss type</td>
<td>Set-up request rejection</td>
<td>Loss of burst</td>
<td>Loss of packet</td>
<td>Loss of packet (BE part)</td>
</tr>
<tr>
<td>Control</td>
<td>Out-of-band</td>
<td>In-band or Out-of-band</td>
<td>In-band</td>
<td>Out-of-band (GST part) In-band (BE part)</td>
</tr>
<tr>
<td>Latency mainly given by</td>
<td>- 3 x prop. delay</td>
<td>- Prop. delay + burst assembly delay, offset</td>
<td>- Prop. delay</td>
<td>- Prop. delay</td>
</tr>
<tr>
<td>Control overhead given by</td>
<td>Connection set-up</td>
<td>BLP</td>
<td>Packet header</td>
<td>Packet header (BE part) and Connection Set-up part if D-WRON for GST</td>
</tr>
</tbody>
</table>

1) Latency is end-to-end delay. This row is valid in the case of no set-up rejections, packet or burst loss.
A main motivation for OBS is to reduce the optical technology complexity. Therefore, avoidance of burst alignment is essential, and asynchronous operation required. Burst assembly mechanism typically includes timers to prevent excessive delay, and variable burst length gives the assembler more freedom to optimise burst lengths. Furthermore, in asynchronous operation, fixed burst length does not in general reduce blocking probability. Hence, we consider in this study that OBS networks operate asynchronously, with variable length bursts, as is the custom in most work on OBS.

On the other hand, as further discussed in [7] OPS has two schemes that in general are more attractive, namely synchronous operation with FLP and asynchronous operation with VLPs. The two schemes’ main advantages are summarised in Table 2.

**Contention resolution**

In the hybrid scheme, contention resolution is not needed for the GST packets since they have an absolute guarantee and follow reserved wavelength paths through the network. For the BE traffic, requirements to contention resolution will be as for OPS or OBS, depending on the scheme chosen.

**Time domain contention resolution**

Buffering enables so-called store-and-forward networks, in which contending packets are stored until they can be forwarded. For unlimited buffer capacities, the network can guarantee that all packets reach their destination. With optimum routing schemes, this scheme also guarantees that the packet follows the shortest possible path, which maximises the network overall capacity.

**Optical buffering**

Most optical buffers proposed rely on Fibre Delay Lines (FDL), which are known to be bulky and to

![Figure 3 Potential and viable (dark red boxes) OBS and OPS packet/burst handling schemes. As mentioned earlier, control information in OBS may also co-propagate in-band with the burst](image)

**Table 2 Main properties of attractive OPS packet handling scheme**

<table>
<thead>
<tr>
<th>OPS packet handling scheme</th>
<th>Synchronous FLP</th>
<th>Asynchronous VLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synchronisers / network clock</td>
<td>Required</td>
<td>Not required</td>
</tr>
<tr>
<td>Edge node assembly</td>
<td>More complex</td>
<td>Less complex</td>
</tr>
<tr>
<td>Architecture required to avoid internal blocking</td>
<td>Reconfigurably non-blocking</td>
<td>Strictly non-blocking</td>
</tr>
<tr>
<td>Requirements to contention resolution</td>
<td>Lower</td>
<td>Higher</td>
</tr>
<tr>
<td>Overhead</td>
<td>Higher</td>
<td>Lower</td>
</tr>
</tbody>
</table>
require stable external temperature to maintain a stable propagation delay. The time spent in the FDL is determined by the fibre length – it is not a RAM. More complex buffer designs, achieved by interconnecting re-circulating FDLs or by broadcast-and-select architectures [14] can increase the accessibility after insertion. However, both the increased loss, which must be compensated for by EDFAs, and the switch cross-talk introduced, decrease signal quality. Variable length packets further increase buffer size and buffer design complexity. Complex FDL multi-stage buffers may be required, and available size in OPS nodes is in practice limited. Alternatively, [15] has demonstrated a variable optical delay circuit for use in optical buffers, suitable for variable length packets. However, this requires extensive hardware such as re-circulating loop, filters, circulators and λ-converters.

Electronic buffering
Most OPS designs assume an electronic control unit, since electronic processing is superior to optical processing. Similarly, electronic memory is inexpensive compared to FDLs. However, electronic buffering requires opto-electronic interfaces. Avoiding such interfaces is one of the main targets in OPS. We believe that buffering in OPS core routers should be used in combination with other contention resolution schemes. Therefore, only a small fraction of the packets undergo buffering, thus only a few buffer input interfaces are needed. An OPS design partly based on electronic buffering is proposed in [16], and a design that relies on electronic buffering only is proposed in [17].

Wavelength dimension contention resolution
By sharing the total traffic on several wavelengths, i.e. decreasing the load on each wavelength, the contention problem is minimized, and with a sufficient number of wavelengths a sufficient packet loss ratio can be achieved [18]. However, because efficient utilisation of the link-capacity is important for achieving cost-effective transport networks, a relatively high load on each of the wavelengths is desirable. Contention can be resolved by using λ-converters to convert packets to vacant wavelengths at the same fibre without increasing the number of wavelengths. Due to properties of statistical multiplexing, this method improves with the number of wavelengths per fibre. Furthermore, the signal quality aspects are potentially excellent, since the non-linear response of some wavelength converters has a re-shaping effect (2R). Adding a clock-recovery and a pulsed local source, full regeneration (3R) can be achieved [19], and the integration potential opens up for compact devices. This must be exploited to achieve cost-effective contention resolution due to the high number of wavelength converters needed in this approach. The tuning speed is an important parameter for tuneable wavelength converters since it determines the optical guard band needed between subsequent packets.

Contention resolution by deflection routing
When two packets contend for the same output fibre/wavelength pair, the conflict can be solved by simply forwarding one of the packets on another fibre, and let an alternative OPS node be responsible for forwarding and switching the packet towards its destination. This contention resolution scheme is called “deflection routing”, and can be seen as spatial contention resolution. In a sense the whole network is used as a buffer. The great advantage of this approach is its simplicity, since buffers and wavelength converters may be avoided. An important drawback is the significant loss of packet order that follows, which increases the edge node complexity. Furthermore, the overall network capacity decreases with the increased average hop-count, so that deflection routing does not work well for high loads, especially in asynchronous networks [20].

Comparison of contention resolution schemes
The main advantages of electronic buffering are compatibility with asynchronous scheme, low footprint, excellent signal quality aspects and the commercial availability. The main drawback is the limited transparency and the potentially higher cost of the O/E interfaces.

Wavelength conversion has the same advantages, and somewhat better transparency properties. Furthermore, the performance improves with the number of wavelengths. Demonstrations of wavelength conversion at high bit rates prove their feasibility.

Optical buffering is transparent, but has no random access in the time domain. FDLs have negative impact on the signal quality. Asynchronous schemes require more advanced scheduling processing than slotted schemes. Relying purely on FDLs in this scheme seems unrealistic due to the large footprint required for extensive FDL arrays or use of complex multi-stage buffers.

Deflection routing’s greatest advantage is the low hardware complexity in the core nodes and its transparency. However, the increased hop count, which limits network capacity, especially for asynchronous networks, is a drawback. Furthermore, deflection routing in slotted networks requires complex edge nodes, since the need for fragmentation in combination with the possibility of loss of packets, will require larger edge buffers and mechanisms capable of reassembling fragments and suitable time-out values to decide when a fragment is lost or simply delayed.
Combining contention resolution and packet handling schemes

A combination of contention resolution schemes should be used for optimum performance/complexity ratio. [21] demonstrates the technical feasibility by implementing contention resolution in an OPS node using wavelength conversion, optical buffering and deflection routing. Wavelength conversion has good signal quality aspects and is effective in resolving contention. To achieve very low PLRs it should be used with buffering, e.g. FDLs as in [22, 23]. Here, the wavelength conversion minimises the buffering, and in slotted scheme a small amount of single unit length FDLs can be used. This combination will not work as well for asynchronous packet handling scheme, which requires more buffers and more complex FDL buffer design. We believe that asynchronous mode will benefit from the random access and simple design of electronic buffers in combination with wavelength conversion.

Node design in OPS

A scalable packet switch design

A modular and scalable design, scaling to a very high number of wavelengths, and a high node degree is shown in Figure 4. The design is based on using \( \lambda \)-converters and Array Waveguide Gratings (AWG), thereby avoiding the use of optical switches. Benefits of this design are that the signal path does not introduce large attenuation, contributing to the scalability of the switch, and that no additional \( \lambda \)-converters are needed when using the wavelength dimension for contention resolution. When a signal is sent into an AWG, the wavelength of the signal will decide at which output on the AWG the signal will occur. This principle is exploited for routing the packets to the desired output. The input WDM signal of each fibre is demultiplexed to its corresponding wavelengths and fed to the input of the \( \lambda \)-converters. The outputs of each \( \lambda \)-converter are then fed to the AWG inputs. By tuning the output wavelength, packets can be sent to any of the AWG outputs. The packet will be sent to the scheduled output if a vacant wavelength can be found. If no output with correct destination is available, the packet will be sent to one of the buffer inputs, if a vacant buffer input can be found. If not, the packet will be dropped. Buffered packets are clocked out of the buffer and sent back to an AWG input as soon as a wavelength output to the destination becomes available. At the buffer output, the wavelength, and thus the output of the first AWG, is decided by tuning a tuneable laser. This type of architecture is called a feedback design, and has the benefit of supporting packet priority also when FDLs are used for buffering [24].

In this design, each of the input fibres is coupled to a corresponding input plane. Hence, the AWG size in each plane equals the sum of the number of wavelengths, ‘W’, and buffer inputs, ‘B’, which is independent of the number of fibres and node degree. Each plane has ‘W’ outputs. The outputs are fed to a second set of ‘W’ AWGs with size N X N, making the maximum allowable number of fibres depend on the AWG size.

The buffer is shared among all the planes. Total number of buffer inputs corresponds to the number of buffer output ports in each plane. The buffer ports from each of the planes are passively coupled together using 1 : N couplers so that port ‘1’ in plane ‘1’ are coupled to port ‘1’ in plane ‘2’ through ‘N’, and the same goes for the rest of the ‘B’ ports. At each of the buffer outputs, wavelengths are set using tuneable lasers. By setting the wavelength, the packet is forwarded to the decided output.
Forwarding of a packet undergoes two stages. In the first stage, the packet’s wavelength at the output fibre is decided. In the second stage, the packet is forwarded to the destined fibre. A packet occurring at one of the input fibres is forwarded to its destination as follows: In each plane, which is the first stage in the forwarding, the wavelengths in each fibre are demultiplexed to \(W\) corresponding fibre-lines. Each line is then fed to a \(\lambda\)-converter, setting the wavelength so that the packet is forwarded to any of the desired outputs of the AWG, making it possible to freely choose at which AWG at the second stage the packet will occur. This makes the first stage wide-sense non-blocking as described for the basic design.

After passing the first AWG, the packet is sent through a \(\lambda\)-converter before occurring at the input of the chosen AWG in the second stage. By setting the output wavelength of the \(\lambda\)-converter, forwarding to any output fibre can be chosen, thereby making the whole switch wide-sense non-blocking. The packet then undergoes a third wavelength conversion, converting to a vacant wavelength at the output fibre before the packet is sent through a passive coupler on to the output fibre.

**Hybrid node design**

For separating SM and GST packets in the node, the State of Polarization (SOP) may be used [9]. Figure 5 illustrates the node design with polarization separation and combination mechanisms as well as an OPS with an electronic packet switch buffer [10]. Polarization BeamSplitters (PBSs) physically separate SM and GST packets at the input interface. Neither guard band between GST and SM packets nor headers on GST packets and processing in the detecting node, are required. Because of polarization variations in the transmission fiber, Automatic Polarization Control (APC) at the node inputs is necessary. It will however operate on a relatively slow timescale of milliseconds to seconds, corresponding to the frequency of variations in the fibre’s physical environments [25]. The viability of a transmission path through a model-node is experimentally verified in [9]. The physical viability of a node input interface, with combined wavelength conversion and header payload separation is proposed and experimentally verified in [26]. An alternative to segregating by polarization is to replace the PBS by a fast switch and send the class information in a packet header [10].

The cross-connect in the node may be static (S-WRON) or dynamic (D-WRON). An S-WRON cross-connect is simple and reliable and does not need remote control, avoiding a control plane. Secondly, in S-WRONs, path setup delay is avoided enabling a GST without waiting time. For D-WRONs, however, connections can be established dynamically for relatively short time intervals, supporting traffic patterns that vary with a moderate frequency.

*Figure 4 The scalable packet switch design, with two switching stages, using a fully shared buffer. In the circle is shown the principle of the second switching stage.*
The aim of our simulations is to compare the performance of synchronous and asynchronous packet scheduling and fixed/variable length packet schemes. Furthermore, we analyse the basic characteristics of a hybrid node. We do this by characterising packet delay and packet loss. In our simulations we assume an optical packet switch using the wavelength dimension for contention resolution in combination with a minimised number of electronic buffer inputs, as described above and in [17]. Since the wavelength dimension can be used for contention resolution, the switch must be capable of doing wavelength-conversion. In this switch, full wavelength conversion is assumed, thus packets arriving at any input with any wavelength can be switched to any output and converted to any wavelength.

Since electronic buffering provides random access in the time domain, packets that are buffered are allowed clocking in and out of the buffer at random times. The switching matrix is assumed to be strictly non-blocking, thus it avoids internal blocking.

Simulation parameters

We assume an optical switch with ‘N’ fibre inputs, each with ‘W’ wavelengths. We have used N * W independent traffic generators, generating fixed or variable length packets according to a Poisson arrival process, corresponding to a load of 0.8 normalised load. The traffic generators generate packets asynchronously. When simulating synchronous packet arrivals, packets are synchronised to discrete time slots at the input, thereby causing a slotted operation of the packet switch.

Both the input and the output traffic are uniformly distributed at the inputs and outputs, respectively. Since electronic buffering is assumed, the packets can stay in the buffer for an arbitrary period of time. The buffer size is set to 100 kB, which in our simulations is sufficient to avoid additional packet loss due to buffer overflow. Packets in the buffer are handled with FIFO priority. We expect that the future transport network will employ links with high channel (wavelength) counts, but have approximately the same node degree as we see in the networks of today. Therefore we have chosen to evaluate the performance of the switch in a backbone network with a
node-degree of ‘8’ and a high link channel count, starting at 32. Performance is analysed varying both number of buffer interfaces and channel count. In the case of asynchronous operation, both fixed and variable packet length is evaluated. The fixed packet length is 500 bytes, while the variable packet length distribution is based on measurements of packet lengths in Internet [4]. It has the following distribution, expressed in number of packet occurrences: 40–44 bytes 62 %, 45–552 13%, 553–576 8 %, 577–1500 17 %.

In [24] the mean delay is shown to decrease quickly as the number of wavelengths increases. We find the same tendency. The longest mean delay observed is with 32 link channels, asynchronous packet arrival and fixed packet length, where a delay as low as $6.8 \cdot 10^4 (\pm 7 \cdot 10^5)$ of the duration of a packet is found. With moderate packet lengths of 500 bytes, and a bit rate of 10 Gbit/s, this corresponds to a 270 ps mean delay. Transmission delay is therefore dominant and will be increasingly dominant as the link channel count increases. Packet loss is therefore the main topic in this analysis.

**PLR performance electronic buffering:**

**Fixed versus variable packet length, asynchronous operation**

In our simulations, we have chosen to vary the number of buffer interfaces. For a given link channel count, this is the parameter to minimise since it is an important cost factor that directly governs the PLR. In Figure 6 we have compared the performance between using fixed and variable length packets when switching asynchronously, assuming an electronic Random Access in Time Memory (RATM) buffer. Results when varying both the number of wavelengths and the number of buffer interfaces are shown.

Results for both fixed and variable length packets (VLP) are shown for 32 and 256 wavelengths. For both wavelength counts there is a marginal performance difference between performance of fixed and variable packet lengths. The difference is close to or within the confidence interval. Since there is no synchronization mechanism at the inputs of the switch, packets do not arrive in distinct timeslots, causing little or no difference in performance when switching fixed or variable length packets. Since the performance difference is small, we see it as beneficial to use VLPs, and therefore for the rest of this paper VLPs are assumed when switching is asynchronous.

As expected, packet loss is shown to decrease as the number of channels is increasing when keeping the number of buffer inputs at a constant level. This is due to the wavelength dimension’s impact on contention resolution. This implies that as the number of channels increases, the number of buffer inputs can be decreased while maintaining an acceptable packet loss probability. When the number of link-wavelengths reaches 256, 15 buffer inputs are sufficient to reach a packet loss probability better than $10^{-6}$. At 256 wavelengths, a PLR of $3.2 \cdot 10^{-5}$ is achieved without using buffers, hence with a sufficiently high number of wavelengths, an acceptable PLR can be reached without using buffering. This is confirmed in Figure 7, where results from simulations without buffering are plotted. We see clearly that the slotted scheme achieves a better PLR. We also observe that when the link channel count increases, the difference in packet loss performance between the two schemes increases. At the most demanding PLR threshold of $10^{-6}$, buffering very few packets is necessary. At 128 channels per link, the highest packet loss probability is that of the asynchronous scheme, with a PLR of $10^{-3}$. This implies that for this high channel count system, even this scheme will only need very limited buffering resources.

In Figure 8, the PLRs for the slotted and asynchronous scheme, for link channel counts of 32 and 128, are shown. When having 32 wavelengths in each link, a PLR of $10^{-6}$ can be achieved by using approximately 50 and 24 buffer interfaces in the asynchronous and slotted schemes, respectively. Thus, the slotted scheme requires less than half as many buffer inputs as the asynchronous scheme. Assuming 128 wavelengths in each link, these numbers decrease to 24 and 14, respectively. The drastic decrease in required buffer inputs, especially for the asynchronous scheme, indicates that the asynchronous approach becomes increasingly attractive with the increasing number of link wavelengths.

![Figure 6 Packet loss ratio (Y-axis) as a function of number of buffer inputs (X-axis). Performance using fixed (broken lines), and variable length packets (continuous lines) are compared. 'W' is the number of wavelengths in each fibre, 'F' indicates fixed packet length. The error bars mark the limits of a 95 % confidence interval. Where only the upper limit is given, the lower limit is lacking. Higher precision can be achieved, making simulation time excessively long.](image)
The GST class is given absolute priority so that its quality is given solely from physical parameters in the WRON network, like transmission delay and signal quality. The quality of the SM (= BE) traffic does however depend on characteristics of the GST traffic, especially on the GST traffic share, which may typically vary. To illustrate how the GST traffic share in an OpMiGua node influences the performance of the SM traffic, we have simulated Packet Loss Ratio (PLR) and buffered delay for the SM traffic. We assume the node design shown in Figure 5 with a passive OXC and an OPS consisting of a strictly non-blocking switching matrix and a limited number B buffer inputs. Parameters are as for the previous simulations. The number of link wavelengths is $W = 32$. The switch experiences a uniform traffic distribution pattern from both the traffic classes. The GST packets are forwarded to a fixed destination and wavelength, thereby avoiding congestion with other GST packets.

We have chosen to evaluate the performance of the system when the GST packet length is fixed. The GST packet length is set to 100 times the mean SM packet length, making overhead caused by the GST packet’s output reservation negligible. The long fixed-length GST packets correspond to an aggregation of several, e.g. IP-packets in a fixed length container.

We assume the buffering to be randomly accessible in the time domain, i.e. electronic buffering within a short time frame, and potentially optical memory within a longer time frame. Consequently, the packets can stay in the buffer for an arbitrary period of time. However they are scheduled to an output as soon as a wavelength to the destination becomes vacant. The Packet Loss Ratio (PLR) and delay results are plotted in Figure 9.

There are two counteracting effects impacting on the PLR. First, the GST packets are given absolute priority and as the GST traffic share increases, the lower priority SM traffic is increasingly penalized, resulting in an increase in PLR. On the other hand, as the GST traffic share increases, the remaining SM traffic share gets a relatively higher share of buffering resources available, resulting in a decrease in PLR. In Figure 9 we find that for low GST traffic shares of up to 30 %, the first effect is dominating, resulting in an increase in PLR. As the GST traffic share increases, the second effect becomes dominant and the PLR decreases significantly. For 90 % GST, no packet loss was observed for the $1 \cdot 10^8$ packets that were simulated. This indicates that given a specific PLR requirement, packet switch resources may be reduced if GST traffic shares are high.
The GST induced contention on SM packets increases with increasing GST traffic share. Consequently, the delay increases because buffered packets will wait longer before a wavelength becomes vacant. In the delay values in Figure 9, the clocking of the packets into the buffer is not accounted for. This corresponds to a cut-through solution, where packets can be clocked out of the buffer while they are still being clocked in. If such a solution is not available, one unit of delay must be added. The magnitude of the delay values indicates that packet reordering within a wavelength channel will occur. However, the bit-rate of a wavelength channel is typically in the order of 10 Gb/s. If the packets of an application are multiplexed evenly into the wavelength channel, e.g. for a 100 Mb/s application, every 100 packet in the wavelength will belong to the application. With mean delay values below 1.5 mean SM packet duration, packet reordering will therefore be of low probability.

**Conclusion**

There seems to be a convergence of the optical packet switching and optical burst switching concepts when it comes to packet/burst handling and contention resolution. Acceptable loss rates as well as QoS differentiation can be achieved in both concepts. In general, we see OBS as a first step towards a real dynamic network, with OPS as the ultimate goal. The hybrid network combines circuit switching with either packet- or burst switching. This concept allows guaranteed service with no packet loss and fixed delay, throughput efficiency of statistical multiplexing, and bypassing of transit traffic through a WRON network allowing smaller packet switches. A hybrid network can be implemented today, combining electronic packet switches with optical circuit switches. Furthermore it enables a migration path towards the technology found most beneficial for the future, be it a pure OPS network or the hybrid combination including an OPS as part of the hybrid node.

The optical packet switch designs discussed in this paper all rely heavily on wavelength conversion, and also on relatively large AWGs. Hence, for these designs to become attractive it is crucial that the mentioned components become cost effective.

Since the complexity of optical buffering is regarded as the chief argument against optical packet switching, we have compared buffering techniques and evaluated their performance in slotted and asynchronous mode. We conclude that optical packet switches using FDLs is an attractive solution if the number of wavelengths in each link is low and the bitrate at each wavelength is too high for electronic processing/buffering. In this case, the slotted scheme appears to be advantageous, as it reduces the complexity of FDL based buffers.

However, if extensive use of wavelength division multiplexing is preferred, and the bitrate at each wavelength is sufficiently low to be processed electronically, electronic buffering appears advantageous. Only a very small number of electronic buffer interfaces are necessary to obtain sufficient performance, even when using the asynchronous scheme.

For high capacity optical packet switches, capable of handling several Tbit/s, it is our view that the limited increase in buffering required for the asynchronous scheme is less complex than the slotted scheme’s requirements for network synchronisation, optical packet alignment, and electronic edge reassembly.

Therefore, the most viable optical packet switched network design could employ an asynchronous, variable packet length handling scheme, with packet switches using a combination of wavelength conversion and electronic random access memory for contention resolution. For even higher link wavelength count, we expect that buffering will not be required, and the network will become transparent.

**Acknowledgments**

Thanks to Telenor R&D and the Research Council of Norway (NFR) for financial support for this work.

**References**


Steinar Bjørnstad received his MSc in electronics from the Institute of Physics, University of Oslo in 1991, and his PhD in telecommunication from the Institute of Physical Electronics at the Norwegian University of Science and Technology (NTNU) in 2004. He worked with electronic design at Ring-Master 1991–1995 and at NERA Contec Design 1995–1996. He was then with Telenor R&D as Research Scientist on optical communication until 2005, when he started work as Associate Professor at the Institute of Telematics at NTNU, and as Senior System Architect on transport networks at Canal Digital. His interests include optical communication systems and networks, systems for television broadcasting, quality of service aspects of networking, discrete event modelling, electronic design and electromagnetic compatibility. He has managed several research projects and been active in several projects funded by the EU within the ACTS, IST and COST programs. He has authored and co-authored a number of papers, and has currently three patents pending.

email: steinar.bjornstad@iet.ntnu.no

Dag-Roar Hjelme is Chief Scientist at Optomed and Adjunct Professor at the Department of Electronics and Telecommunication, NTNU, both Trondheim. He earned his Siv.Ing. degree in Electrical Engineering from the Norwegian Institute of Technology (NTH), Trondheim, and his PhD in Electrical Engineering from the University of Colorado, USA. Prior to joining Optomed he was Senior Research Scientist at the Department of Photonics, SINTEF Electronics and Cybernetics, where he managed research projects on optical sensor technology and optical communication systems. After receiving his PhD he was Post-Doctoral Research Associate and later Staff Scientist at the University of Colorado, where his research focused on semiconductor lasers, integrated optics, microwave optics, and fiber optic systems. He is the author and co-author of more than 30 refereed journal papers and 70 conference presentations. He is co-inventor of six international patents. Dr. Hjelme is a member of the Norwegian Academy of Technological Sciences.

email: Dag.R.Hjelme@optomed.no

For a presentation of Norvald Stol, please turn to page 146.