Service Differentiation in Optical Burst Switching Networks*

Christoph Gauger, Klaus Dolzer, Jan Späth**, Stefan Bodamer
University of Stuttgart, Institute of Communication Networks and Computer Engineering (IND),
Pfaffenwaldring 47, 70569 Stuttgart, Germany
e-mail: {gauger, dolzer, spaeth, bodamer}@ind.uni-stuttgart.de

Abstract

In this paper, we give an overview and classification of optical burst switching schemes and burst reservation concepts. We compare the performance of different burst reservation mechanisms for an OBS node that does not distinguish different classes. The one that performed best and allows service differentiation, called Just-Enough-Time, is then evaluated by simulations and an approximative analysis for a two-class OBS node. A variety of new results show the pros and cons of the evaluated reservation mechanism with respect to service differentiation.

1 Introduction

At the beginning of the new millenium several trends can be observed in the field of communications networks. First, bandwidth requirement in networks seems to grow without limits. IP (Internet Protocol) based data networks play a central role. This is not only due to the fact that data traffic has surpassed voice traffic but even more due to the exponential growth rate of IP traffic volumes. Second, more and more users and applications request QoS (Quality of Service) mechanisms from today's communication networks. Third, optical technology continues to provide an exponential growth at higher rate than IP traffic growth in fiber transmission capacities.

In this paper, we will elaborate on these trends and show how they motivate *optical burst switching* (OBS) as a new switching paradigm for future transport networks. We start with an overview of the evolution of photonic and IP networks, and classify OBS with respect to switching paradigms. Section 2 surveys the definition and design parameters of OBS. In Section 3, different proposed reservation mechanisms are introduced and compared in a scenario where no classes are distinguished. Section 4 describes an analysis for the burst loss probabilities with JET for arbitrary offset values. Finally, in Section 5 a detailed discussion of the performance in a two-class OBS node is presented.

1.1 Evolution in Photonic networks

In the late 70s, the first fiber based optical transmission systems were installed. Today, most wide area traffic in communications networks is carried via fibers. Until a few years ago, most systems used a single high-speed optical channel and all multiplexing was done in the electrical domain. In 1995, a new technology entered the market in the USA: *wavelength division multiplexing* (WDM) [20]. This optical multiplexing technique allows better exploration of fiber capacity by simultaneously transmitting multiple high-speed channels on different frequencies (wavelengths) [14, 19, 23].

Fig. 1 shows a possible evolution scenario for photonic networks based on WDM. WDM is still mainly used on point-to-point transport links. Today, add/drop multiplexers (ADM) and cross-connects (CC) become available. ADMs and CCs allow the realisation of ring and mesh networks, respectively. For the future, there is a clear trend towards higher reconfiguration speeds in these networks [21]. In the long term, optical packet switching seems to be a promising technology. However, due to its complexity optical packet switching is expected to remain a research topic for some more years to come.

Recently, OBS was proposed as a new switching paradigm for optical networks requiring less complex technology than packet switching. OBS is based on some concepts developed several years ago for electronic burst switching networks. At that time, burst switching essentially was an extension of fast packet switching with packets of variable and arbitrary length employing decentralized shared buffer switches [1, 2, 15]. OBS has some more specific features and will be described in detail in Section 2.

Another hot topic at the moment is extending Multi Protocol Label Switching (MPLS) concepts [25] to optical transport networks (so-called MPλS) [4, 13]. Originally developed to increase forwarding speed by using short label information, work in the MPLS domain includes more and more traffic engineering and traffic management aspects [3]. For MPλS, the core idea is to use wavelength channels as labels and to establish appropriate routing paths in the network. Such paths allow fast switching of

* This work was partly funded within the TransiNet project (www.tranisnet.de) by the German Bundesministerium für Bildung und Forschung under contract No. 01AK020C.

** Jan Späth is now with Marconi Communications, Backnang, Germany. Email: jan.spaeth@marconi.com

data without requiring complex routing processes along the path. Label switching concepts can be easily integrated with burst switching concepts [18].
Label switching as well as burst switching concepts serve a more efficient integration of IP and WDM than allowed by today's multi protocol stacks. Both concepts can be combined to a comprehensive and efficient "IP over WDM" framework [8, 12, 18, 5].

1.2 Evolution in IP networks

The Internet is a packet oriented network based on IP, a connectionless networking protocol. The Internet has been designed to offer best effort services and for a long time this was sufficient. But recent years have seen an increasing demand for QoS mechanisms mainly due to new applications, an increasing number of users and traffic volume, and growing commercial interest in network services.
On the one hand, this lead to the development of new network technologies like ATM (Asynchronous Transfer Mode) which allow a broad spectrum of service guarantees. On the other hand, there is significant effort to include QoS mechanisms into the Internet. These mechanisms can be classified as providing either *absolute* or *relative* guarantees represented by IntServ [7] and DiffServ [6] approaches, respectively.
It is a key feature of the Internet that it can be run basically on top of any transport technology. This independence of the physical layer strongly contributed to the wide spread use of Internet technologies. Nowadays, Internet traffic is the dominant part in many networks. Therefore, more and more networks are designed in an "IP centric" way. This includes a transport layer offering most efficient support for IP traffic. OBS is one proposal of how to realise such a transport network

1.3 Switching paradigms

The basic switching concepts are circuit switching and packet switching. For their application in an optical transport network, their pros and cons can be characterised as follows.
Circuit switching (of wavelength channels) is relatively simple to realise but requires a certain amount of time for channel establishment and release independent of the connection holding time. This overhead, which is mainly determined by the end-to-end signalling time, leads to a poor channel usage if connection holding times are very short. For long holding times, circuit switching is very efficient from a signalling overhead point of view. However, that case leads to a reduced ability to adapt to traffic dynamics. This is especially true if IP traffic with its bursty behaviour [9] is carried on top of such a circuit switched wavelength network.

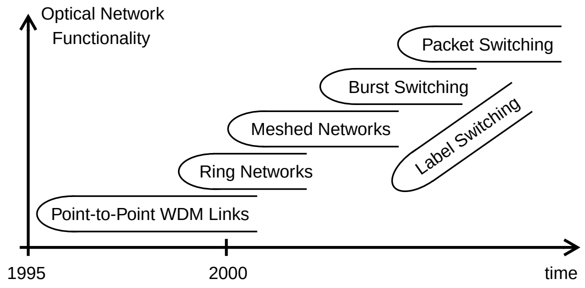

Fig. 1 Evolution of photonic transport networks

Packet switching in the optical domain allows a good adaptation to the dynamics of any higher layer. However, there are several other drawbacks. The first is concerned with realisation aspects. If the realisation is based on opto/electrical conversion, it can be done with technology available today. But this approach suffers from the fact that the development of electronics cannot keep pace with the rapid growth of optical transmission speed. This could be improved by all-optical packet switching technology (including signal processing). Such all-optical approaches will be difficult to realise in the foreseeable future e.g. due to highly complex technology and lack of optical buffers.
Another basic restriction stems from the fact that packets have to be of limited size due to several reasons (buffering requirement in each node, increasing delay if store-and-forward is used). Moreover, each switching process needs a finite non-zero time. This leads to reduced efficiency for large data blocks which have to be transmitted using multiple packets..
As a new paradigm, OBS tries to combine the advantages of both, circuit and packet switching while avoiding the main drawbacks described above.

2 Optical Burst Switching

2.1 Definition and motivation

As mentioned above, OBS is in some way a combination of optical packet and circuit switching. Although there is no unique definition of optical burst switching in literature, it is widely agreed that the following list describes its main characteristics.

- OBS granularity is between circuit and packet switching.
- There is a separation between control information (header) and data. Header and data are usually carried on different channels with a strong separation in time (see example OBS network link in **Fig. 2**).

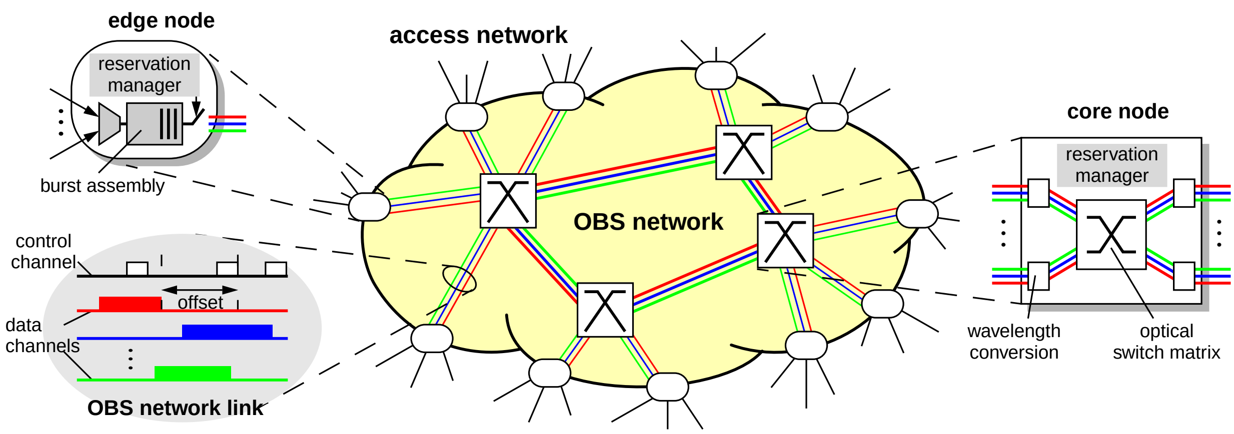

Fig. 2 Node and network architecture for optical burst switching

- Resources are allocated without explicit two-way end-to-end signalling, instead so-called one-pass reservation is applied.
- Bursts may have variable lengths.
- Burst switching does not require buffering.

Note that not all of these features must be satisfied and "smooth" transitions to packet and to (fast) circuit switching are possible.

Although the concept of burst switching has been already known since the 1980s, it has never been a big success in electrical networks. The main reason is that its complexity and realisation requirements are comparable to that of more flexible electronic packet switching techniques (like e.g. ATM).

However, with the introduction of very high speed optical transmission techniques this has changed. Now, there is an even increasing discrepancy between optical transmission speed and electronic switching capability. Moreover, due to cost and complexity aspects, it is advantageous to keep data in the optical domain and to avoid opto/electronic conversion. On the other hand, all-optical packet switching is still too complex to perform all processing in the optical domain.

Therefore, a hybrid approach like burst switching seems very promising: it keeps data in the optical domain but separates control information which allows sophisticated electronic processing of this control data. **Fig. 2** shows some of the main characteristics of an OBS network. There are two types of nodes. In edge nodes, traffic is collected from access networks and assembled into larger data units, so-called bursts. Core nodes serve as transit nodes in the core network. Their main task is switching bursts without extensive processing. To achieve this, some control information containing reservation requests is necessary ahead of every burst's transmission time.

There are several possibilities how to perform reservation of data channel bandwidth. Our paper concentrates on the evaluation of so-called SCDT (*separate control, delayed transmission*) schemes. These reservation concepts are based on a strong separation of control information and data. A reservation request is sent in a separate control packet on a different channel while the actual transmission of the data burst is delayed by a certain basic offset (see **Fig. 2**). This basic offset enables the intermediate nodes to process control information and set up the switching matrix. In contrast to systems with immediate transmission*, which send control information together with the burst, the network can do without buffering the data burst in each node along the path. SCDT, however, requires higher complexity in edge nodes and introduces additional delay to bursts. The basic offset has to compensate for the sum of processing times in all intermediate nodes. Therefore, some upper limit of the number of intermediate nodes has to be known prior to reservation which requires some kind of source routing. In each core node, offset information in the header has to be reduced by the actual processing delay.

SCDT schemes use one-pass reservation, i.e. the sender of a burst does not wait for an acknowledgement of its reservation request. This approach is in contrast to two-pass reservation as typically applied during connection setup in circuit switched optical networks. The advantage of a one-pass reservation is higher efficiency as there is no overhead caused by propagation delay. An example may illustrate this. The transmission time of a 100 kbyte burst on a 10 Gbit/s link is 80 µs while the propagation delay over a distance of 200 km (which is not long in a backbone network) is typically about 1 ms.

*Non-SCDT schemes with data immediately attached to control information could be imagined, but are very similar to either fast packet or fast circuit switching.

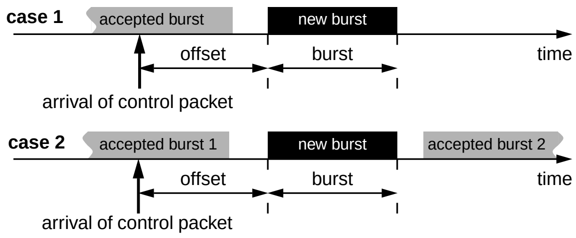

Fig. 3 Reservation scenarios

2.2 OBS design parameters for SCDT schemes

The following list describes the most important design parameters for OBS and includes examples from literature..

- *Buffers for data bursts at intermediate nodes*. Many proposals avoid buffers or use only simple delay lines to keep the system significantly less complex than a packet switching system [16, 24, 26], other work includes sophisticated buffering concepts [28].
- *Resource reservation mechanism*. Key system resources which have to be reserved are channels and possibly buffers. There are several proposals in literature which are classified and compared in Section 3.
- *QoS support*. First proposals for burst switching only considered one class of bursts [24, 26]. Due to the increasing importance of QoS support, recent proposals extended the OBS concept to multiple service classes [27, 28].
- *Protocol aspects*. Designing a protocol for OBS strongly depends on the reservation mechanism and QoS support to be realised but still offers many degrees of freedom. Even for the one-pass reservation scheme we focus on, “one-way” [16] or “two-way” [26] protocols are possible. In the latter case, blocking events or successful channel reservations are reported back. Note that even with two-way protocols in an SCDT scheme burst transmission starts before any confirmation message is received at the initiating node.
- *Node architecture and technology.* Depending on the design choices for the parameters listed above, there are many realisation possibilities for a burst switching node. Basic building blocks are I/O interfaces, control information processing units such as a reservation manager, and switching systems for control and user data possibly including buffers (see **Fig. 2**). [24] gives a very detailed description of an example node architecture, [28] describes various delay line concepts.
- *WDM technology.* All OBS proposals using WDM as transmission technology require full wavelength conversion in a core node such that each burst can be switched to any of the output channels. Therefore, there is a trade-off between performance benefits due to higher number of wavelength channels and higher cost due to more wavelength converters [21, 22].

3 Comparison of Reservation concepts

3.1 Reservation mechanisms

Recently, several SCDT-based reservation mechanisms have been proposed. They can be distinguished based on their way of indicating the end of a burst and the time when allocation of a WDM channel starts.

A rather simple approach is to indicate the end of a burst by an additional trailing control packet* or using an *in-band terminator (IBT)*. In both cases, there is no information about burst length when the heading control packet containing the reservation request arrives. A mechanism that follows that principle is *just-in-time* (JIT) reservation [26]. Upon arrival of the reservation request a wavelength channel is immediately allocated if available. Otherwise, the request is rejected and the corresponding data burst is discarded. The wavelength channel remains allocated until burst transmission has finished. The only information that has to be kept record of in network nodes is whether a wavelength channel is currently available or not. This makes JIT a light weight approach with low complexity in both edge and core nodes. The drawback of JIT is, however, its reduced efficiency as losses also occur in cases without any transmission conflict between different bursts on the same wavelength (case 1 in **Fig. 3**)

An improvement to schemes like JIT can be achieved by using RLD (*reserve-a-limited-duration*). Mechanisms based on RLD require the sender to signal the burst length in the control packet. A wavelength channel is only allocated for a limited duration so that subsequent burst transmission requests with a start time greater than the finishing time of an allocated burst may be accepted (case 1 in **Fig. 3**). That means the basic offset interval of a burst may overlap the transmission phase of a previously accepted burst. Thus with an IBT approach, the new burst (case 1 in **Fig. 3**) is lost because at the instant when the control packet arrives, the end of the previous (accepted) burst is unknown. In contrast, the end is known with RLD and hence the new burst can be accepted.

* Qiao and Yoo denote this category as TAG (tell-and-go) mechanisms [17, 18].

The *Horizon* mechanism proposed by Turner in [24] is one representative of RLD-based mechanisms. In Horizon, wavelength channel state information is enhanced by the so-called reservation horizon, i.e. the time until which the wavelength is allocated. When a new request arrives Horizon looks for the wavelength with the largest reservation horizon less than the start time of the new burst. Like in JIT, reservation starts immediately upon arrival of the control packet and lasts until the expected end of burst transmission, which is the new reservation horizon of the corresponding wavelength.

Even higher efficiency may be achieved if start times of burst transmissions are also considered for reservation, i.e. reservation does not begin immediately when a request arrives but is delayed by the basic offset. This approach is called RFD (*reserve-a-fixed-duration*) as the channel is allocated for a fixed duration corresponding to the burst transmission time. One proposal of an RFD-based reservation mechanism is *just-enough-time* (JET) developed by Qiao and Yoo [16, 18]. State information in JET comprises both, the starting and finishing times of all accepted bursts, which makes the system rather complex. On the other hand and in contrast to Horizon, JET is able to detect situations where no transmission conflict occurs although the start time of a new burst is earlier than the finishing time of the already accepted burst 2 (case 2 in **Fig. 3**), i.e. a burst can be transmitted in between two already reserved bursts. Hence, bursts can be accepted with higher probability than in Horizon especially in case of large offset time variation.

Qiao and Yoo take advantage of that property and extend this mechanism in order to support different service classes [27]. In this case, the offset of a data burst consists of a base component (*basic offset*) representing the sum of processing times for the control packet and an extra component (*QoS offset*) specific to a service class. As bursts with larger offsets experience lower blocking larger offset values are assigned to high priority classes. We will come back to this extension later in this paper.

3.2 Model and analysis for a single class

The performance of the different reservation mechanisms presented in the previous section can be expressed in terms of the burst loss probability. If we restrict evaluation to a single node case with fixed offsets δ for all bursts the loss probability may be obtained analytically. In case of JET this means that only a single service class is considered.

Under the assumption that control packets (and in consequence data bursts) arrive in a Poisson stream with rate λ we can use the well-known Erlang's B formula for the loss probability of an M/G/n loss system.

$$B(A, n) = \frac{A^n / n!}{\sum_{i=0}^{n} A^i / i!} \qquad (1)$$

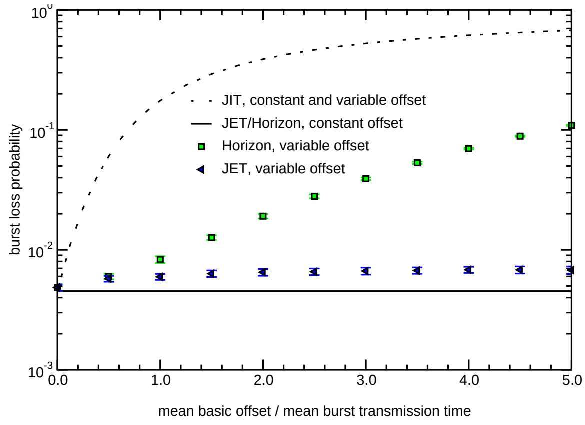

Fig. 4 Dependence of burst loss probability on offset

In this formula n represents the number of servers in a loss system which in this context corresponds to the number of wavelength channels on a link. The offered load A relevant for loss computation depends on the reservation mechanism. For Horizon and JET the offered load is simply the product of arrival rate and mean transmission time h of a data burst. So burst loss probability is given by

$$P_{\text{Loss, Horizon}} = P_{\text{Loss, JET}} = B(\lambda \cdot h, n) . \qquad (2)$$

Note that Horizon and JET have the same performance under the given assumptions as the second scenario shown in **Fig. 3** does not occur in the single node case with constant δ.

If JIT is applied as reservation mechanism the system behaves like a loss system with increased offered load, resulting in the loss probability

$$P_{\text{Loss, JIT}} = B(\lambda \cdot (h + \delta), n) . \qquad (3)$$

The reason for this is that each request blocks a channel for an interval which length is the sum of basic offset and burst transmission time. The increased load leads to a higher loss probability of JIT compared to Horizon and JET especially for large δ as demonstrated by the lines in **Fig. 4** for 16 wavelengths and a total load of 0.5. Therein as well as in several following graphs, we depict the burst loss probabilities against the mean offset standardized by the mean transmission time in order to ease interpretation.

A derived measure especially interesting for dimensioning is the maximum burst arrival rate λ_{max} that can be allowed to achieve a certain loss probability on a link with a given number of wavelengths. From (3) we can conclude that in the case of JIT λ_{max} is reduced by a factor of

$$\frac{\lambda_{max, JIT}}{\lambda_{max, JET}} = \frac{1}{1 + \delta / h} \qquad (4)$$

as compared to Horizon and JET. **Fig. 5** indicates that JIT drastically remains behind JET and Horizon even for relatively small δ. One can see from the figure that a JET/Horizon system with 16 wavelength is even better than a 32 wavelength system using JIT if $\delta > 1.7h$.

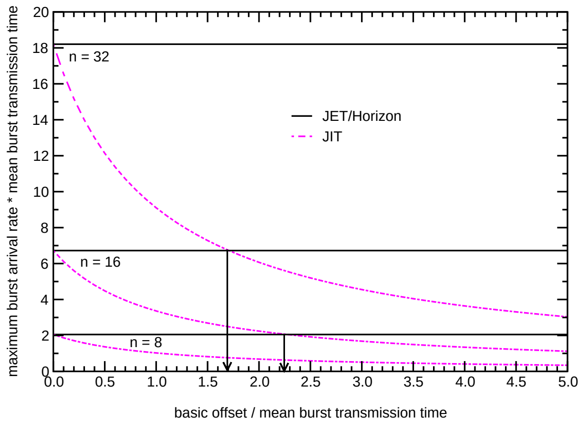

Fig. 5 Maximum burst arrival rate for given loss probability 10^{-3}

In a network scenario, the offset values occurring in a node will not be constant. Therefore, we also investigated the influence of randomly varying δ by simulations as our analysis does not cover varying offsets (**Fig. 4**). For JIT this has no effect, i.e. the loss probability can still be determined using (3). In the case of JET and Horizon, however, we found by simulation that this variation leads to higher losses (variable offset results in **Fig. 4** are obtained for negative-exponentially distributed δ and burst length). While this effect is minor for JET, loss probability significantly increases for a larger mean offset when Horizon is applied. The conclusion is that the higher complexity of JET as compared to Horizon results in better performance for varying offsets.

4 Analysis for two classes

In this section, we present an approximative analysis of the loss probabilities in a JET OBS node that distinguishes two classes. One motivation why a network should support only two classes – stream and elastic – is the debate in the Internet community and recent results indicating that this QoS support might be sufficient [11]. For the following analysis, we assume that class 0 has priority over class 1.

Unlike the single class case where all bursts have the same fixed basic offset to compensate switching and processing times we follow – as mentioned in Section 3.2 – Qiao's and Yoo's suggestion [27] to introduce an additional offset for the high priority class, called *QoS offset*, that provide service class differentiation.

If the base offset and the QoS offset are constant the degree of isolation between the classes solely depends on their effective offset difference, i.e. the constant base offset which is equal for both classes has no impact on isolation. This stems from the fact that this constant base offset δ can be interpreted as a constant shift in time of the reservation process and thus neither arrival nor reservation events are reordered in time. This result has also been proven by simulation for various arrival and service time distributions as well as offset values. Hence, we introduce the effective offset difference $\Delta_{0,1}$ between both classes

$$\Delta_{0,1} = \delta_0 - \delta_1 = (\delta_{\mathrm{QoS},0} + \delta) - \delta = \delta_{\mathrm{QoS},0} . \qquad (5)$$

As the constant base offset has no impact on isolation, we choose the basic offset $\delta = 0$ without loss of generality for all further evaluations.

4.1 Basic formulae

In [27] it has been shown by simulation that the conservation law is satisfied for an OBS system as considered here. If this conservation law holds, the overall loss probability is not dependent e.g. on the number of classes. Thus, the overall burst loss probability $P_{\mathrm{Loss,\,all}}$ in a two-class OBS node can be obtained from Erlang's loss formula (1) in case of Poisson arrivals for an overall offered load A_{all} and bundle size n independent of service differentiation as

$$P_{\mathrm{Loss,\,all}} = B(A_{\mathrm{all}}, n) . \qquad (6)$$

In order to calculate the burst loss probability of the high priority class $P_{\mathrm{Loss},0}$, not only the offered load A_0 of the high priority class has to be considered but also a fraction of the carried traffic of the low priority class. This low priority traffic $Y_1(\Delta_{0,1})$ represents bursts which started transmission prior to the arrival of the high priority control packet and are still being served when the high priority burst starts, i.e. $\Delta_{0,1}$ after the high priority packet arrived. This additional traffic stems from the fact that in this system, high priority traffic is not totally isolated from low priority traffic. Thus, $P_{\mathrm{Loss},0}$ is approximated by

$$P_{\mathrm{Loss},0} = B(A_0 + Y_1(\Delta_{0,1}), n) . \qquad (7)$$

The burst loss probability of the low priority class $P_{\mathrm{Loss},1}$ can be obtained according to the conservation law solving

$$\lambda_{\mathrm{all}} \cdot P_{\mathrm{Loss,all}} = \lambda_0 \cdot P_{\mathrm{Loss,0}} + \lambda_1 \cdot P_{\mathrm{Loss,1}} \qquad (8)$$

with arrival rates λ_{all}, λ_0 and λ_1, respectively. This averaging weights burst loss probabilities with respect to their occurrence.

For the carried traffic $Y_1(\Delta_{0,1})$ we have

$$Y_1(\Delta_{0,1}) = A_1 \cdot (1 - P_{\mathrm{Loss,1}}) \cdot (1 - F_1^f(\Delta_{0,1})) \qquad (9)$$

where $A_1 \cdot (1 - P_{\mathrm{Loss,1}})$ is the carried traffic of the low priority class at the time when the high priority control packet arrives. $1 - F_1^f(\Delta_{0,1})$ is the complementary distribution function of the forward recurrence time of the burst transmission time at time $\Delta_{0,1}$. It describes the probability that a low priority burst that has already started transmission prior to some random observation time τ (the time when the control packet of the high priority burst arrived) has not finished transmission within the period $[\tau, \tau + \Delta_{0,1}]$. (9) is an approximation because in reality longer bursts are discarded with a higher probability [10].

4.2 An iterative solution

According to (7), (8) and (9), there is a mutual dependency between $P_{\text{Loss, 0}}$ and $P_{\text{Loss, 1}}$. Therefore, we suggest an iterative solution for above formulae.
We initialize the iteration with estimates for the loss probabilities of the high and low priority classes, $P^{(0)}_{\text{Loss, 0}}$ and $P^{(0)}_{\text{Loss, 1}}$, respectively. These zero order estimates are given in (10) and can be derived from (6) - (8) by decoupling the high priority class from the low priority class which is equivalent to neglecting $Y_1(\Delta_{0,1})$.

$$\begin{aligned} P^{(0)}_{\text{Loss, 0}} &= B(A_0, n) \\ P^{(0)}_{\text{Loss, 1}} &= \frac{1}{\lambda_1} \cdot (\lambda_{\text{all}} \cdot P_{\text{Loss, all}} - \lambda_0 \cdot P^{(0)}_{\text{Loss, 0}}) \end{aligned} \tag{10}$$

Similar formulae are also published by Qiao and Yoo [27] and yield lower limits for our analysis if the QoS offset is very large (**Fig. 9**, see below).
The distribution function of the forward recurrence time of the burst transmission time is given by

$$F_1^f(t) = \frac{1}{h_1} \cdot \int_{u=0}^{t} (1 - F_1(u))du \tag{11}$$

where h_1 and $F_1(u)$ represent the mean and the distribution of the burst transmission time, respectively. Finally, the amount of carried low priority traffic is determined by (9) using (10) and (11)

$$Y_1^{(0)}(\Delta_{0,1}) = A_1 \cdot (1 - P^{(0)}_{\text{Loss,1}}) \cdot (1 - F_1^f(\Delta_{0,1})) \tag{12}$$

and can be inserted in (7) yielding a first order result for the loss probability $P^{(1)}_{\text{Loss, 0}}$ of the high priority class. By application of (8) and the just derived result for $P^{(1)}_{\text{Loss, 0}}$ a first order result for $P^{(1)}_{\text{Loss, 1}}$ is obtained. Iteration until some precision criterion is satisfied leads to $P_{\text{Loss, 0}}$ and $P_{\text{Loss, 1}}$.

4.3 Special case: negative-exponentially distributed bursts

From formulae given in Section 4.1, it can be seen that $P_{\text{Loss, 0}}$ is dependent on the burst length distribution of the low priority class whereas it is independent of the burst length distribution of its own class. Section 5.2 will elaborate on the impact of low priority burst length characteristics on high priority burst loss probability.
As the negative-exponential distribution has the property that its forward recurrence time is also negative-exponentially distributed, we have from (7) and (9)

$$P^{(i)}_{\text{Loss, 0}} = B(A_0 + A_1 \cdot (1 - P^{(i-1)}_{\text{Loss,1}}) \cdot e^{-\mu\Delta_{0,1}}, n) . \tag{13}$$

From (13), it is obvious that the influence of low priority traffic on $P_{\text{Loss, 0}}$ decreases exponentially for increasing $\Delta_{0,1}$.

5 Performance Evaluation of an OBS node implementing JET

In the following, we use burst with mean burst length of 12500 Bytes with a line rate of 2.5 Gbps per wavelength channel resulting in a mean burst transmission time of 40 µs. Load stands for total load per wavelength comprising high as well as low priority traffic. Unless stated differently, the interarrival time ist negative-exponentially distributed. Guard times for switching are neglected.
In Section 5.1 we keep the offset unchanged and vary the load while in Section 5.2 we keep the load to 0.6 with a high priority load share of 30%and vary the offset.

5.1 Impact of load conditions on service differentiation

Fig. 6 justifies the assumption of Markovian arrivals used in the analysis. Here, $P_{\text{Loss, 0}}$ is depicted for different low priority interarrival time distributions. It can be seen that the high priority class is hardly affected. Therefore, our assumption of negative-exponentially distributed interarrival times yields reasonable results.
In **Fig. 7** the loss probabilities of an OBS node that distinguishes two classes are depicted against the load. Relative traffic shares of both classes are kept constant, i.e. the high priority class is fixed to 30%. In this case, we assume a QoS offset of one mean burst transmission time, i.e. $\Delta_{0,1} = h_1$. Besides the absolute loss probabilities, the loss ratio of both classes is depicted showing that a good grade of isolation is provided over the whole load range, even for high loads. Looking closer at the loss ratio one can see that with increasing load it slightly increases which leads to a changed service differentiation. However, if the focus is

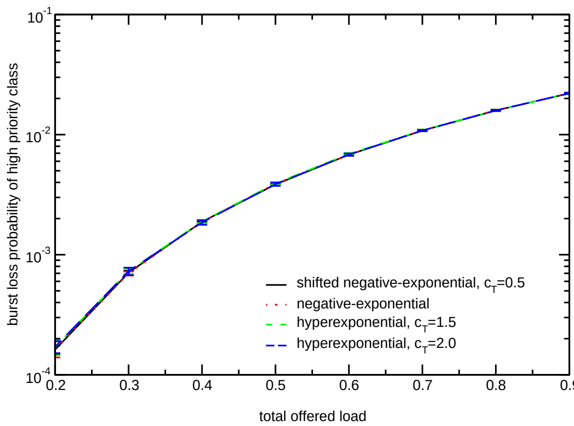

Fig. 6 Different low priority interarrival time distributions ($n = 4$, $\Delta_{0,1}/h_1 = 3$)

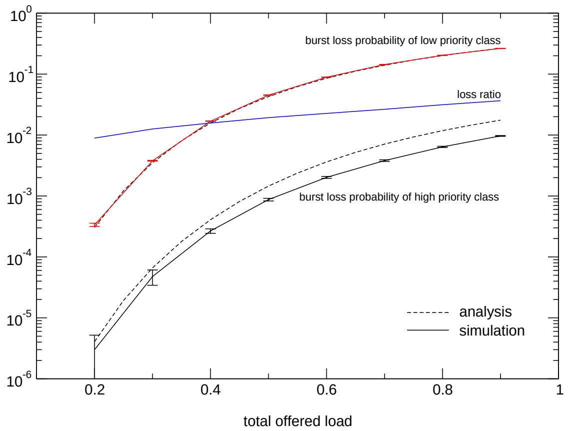

Fig. 7 Analytical and simulation results for loss probabilities and loss ratio against load ($n = 8$, $\Delta_{0,1}/h_1 = 1$)

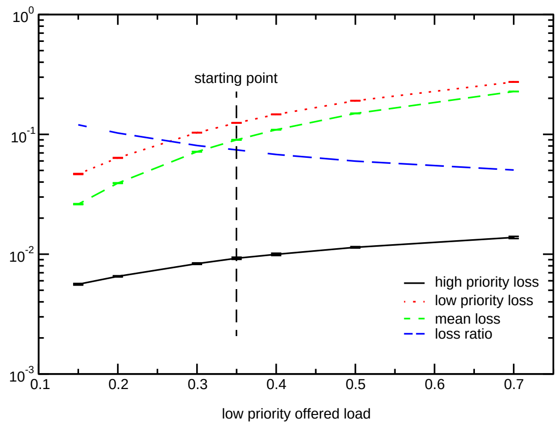

Fig. 8 Varying low priority traffic with constant high priority load of 0.15 ($n = 4$, $\Delta_{0,1}/h_1 = 3$)

on a certain load interval, the JET protocol offers almost constant service differentiation. Curves obtained by analysis and simulation match well for all load values.

Also of interest is the sensibility of the high priority class to load fluctuations of the low priority class. **Fig. 8** shows the grade of isolation of a high priority class from a low priority class with varying load in a scenario with only 4 wavelengths. Here, A_1 – and thus A_{all} – is varied around an initial configuration whereas A_0 is kept constant at a load of 0.15. It can be seen that $P_{Loss,\,1}$ and $P_{Loss,\,all}$ increase significantly whereas $P_{Loss,\,0}$ is only slightly affected. For a higher number of wavelengths the effect on the high priority class diminishes and consequently QoS can be guaranteed almost independent of low priority traffic. If only A_0 is increased, both loss probabilities increase.

5.2 Impact of QoS offset and traffic characteristics on service differentiation

In case of a given traffic with well-defined characteristics such as the overall load, relative load and burst length distribution of each class, the only way to influence differentiation of loss probabilities is to change the QoS offset. In order to determine a reasonable value of $\Delta_{0,\,1}$ for the point of sufficient isolation, following evaluations illustrate the impact of QoS offset in various scenarios.

Fig. 9 shows $P_{Loss,0}$ against the QoS offset normalized by the mean burst transmission time in an OBS node with 8 wavelengths. Results are obtained analytically and by simulation. As bursts are assembled at the edge of the optical network, e.g. by aggregating IP packets, the burst length distribution strongly depends on the aggregation strategy. Hence, we also compare results obtained for different low priority burst length distribution functions (DF). According to Section 4.3, $P_{Loss,0}$ does not depend on high priority burst length DF. **Fig. 9** also includes lower and upper boundaries for $P_{Loss,0}$ referring to the case of neglecting the low priority traffic influence and the case of no isolation, respectively. $P_{Loss,0}$ is depicted for the following low priority burst length DF (with same mean value): negative-exponential DF, uniform DF between 0 and 2 mean burst transmission times and second order hyperexponential* DF with CoV of 2 and 4.

It can be seen that the analysis and simulation match quite well for all DFs. The analysis slightly overestimates the simulation because it assumes a loss probability independent of the burst length. In [10] we have shown that this is not exactly true as longer bursts are discarded with higher probability than shorter bursts. It should be emphasized here that the lower boundary – which is valid for all low priority DFs – is approached very slowly for the scenarios with hyperexponentially distributed low priority bursts which is also indicated by formulae (7) and (9). This fact is not covered by the evaluations presented e.g. in [27] and therefore leads to results which highly differ from the real system behaviour. Thus, for a given traffic and a desired service differentiation, the QoS offset might have to be chosen very large which causes undesirably long delays for the high priority class. In order to avoid this, the aggregation strategy should avoid producing low priority bursts with such unpleasant DFs.

Now, we focus on the low priority traffic for the just described scenario. In **Fig. 10** $P_{Loss,1}$ is depicted for the DFs also presented in **Fig. 9** and yields two important implications. First, it shows that for all DFs the low priority burst loss probability hardly changes for QoS offsets greater than the mean burst transmission time. Conse-

* This hyperexponential distribution satisfies the symmetry condition $p \cdot h_1 = (1 - p) \cdot h_2$ where p is the branch probability and h_1 and h_2 are the mean values of the respective phases.

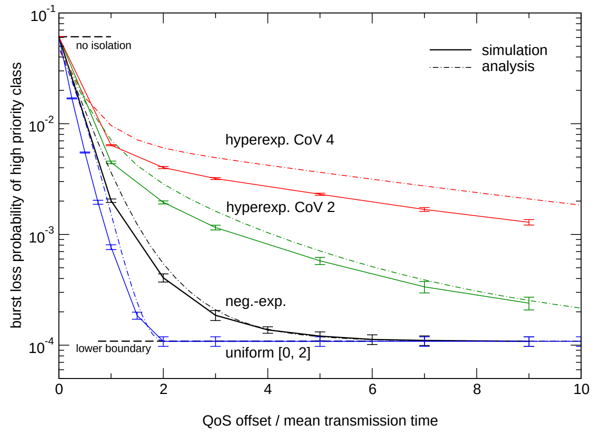

Fig. 9 Comparison of analytical and simulation results for high priority burst loss probability ($n = 8$)

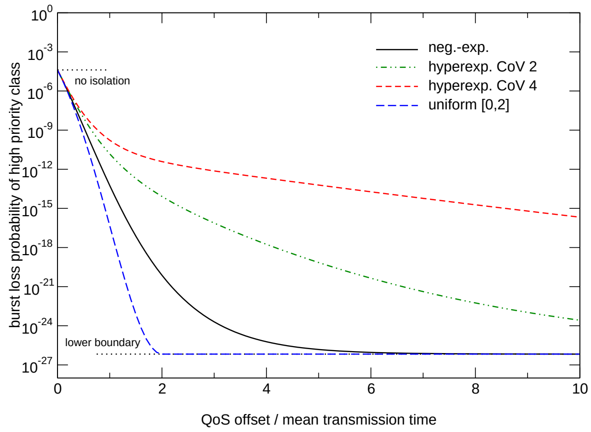

Fig. 11 Analysis of high priority burst loss probability ($n = 64$)

quently, the ratio of $P_{\text{Loss},0}$ and $P_{\text{Loss},1}$ approximately follows the curve of the respective high priority burst loss probability. Second, **Fig. 10** depicts a limit of our approximative analysis. All curves obtained analytically approach the same boundary, whereas the simulated curves do not converge. An explanation is the approximative application of the conservation law described in Section 4.1.

Fig. 11 shows an effect that generalizes the previous discussed behaviour of an JET OBS node. Here, an OBS node with 64 wavelengths is evaluated. It can be observed that the run of the curves matches the one presented in **Fig. 9** whereas the order of magnitude of losses changes significantly.

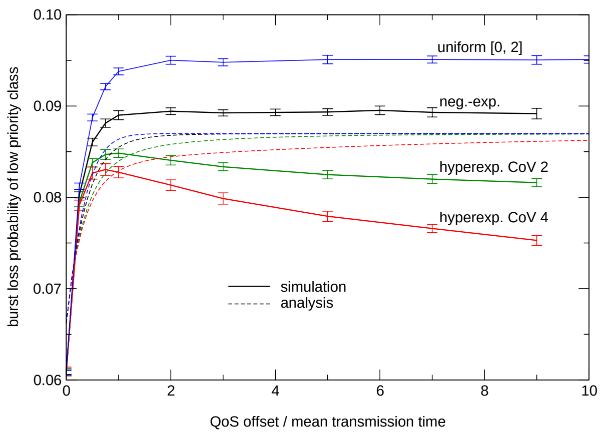

Fig. 10 Comparison of analytical and simulation results for low priority burst loss probability ($n = 8$)

6 Conclusion und future work

Based on a discussion of various switching paradigms as well as photonic and IP network evolution we showed that optical burst switching promises many benefits for future QoS supporting high speed transport networks. We gave a detailed overview of characteristics and design parameters of optical burst switching. Then, a classification of different reservation mechanisms proposed in literature as well as a performance comparison for a single OBS node was presented.

In single-class optical burst switching, we found that JET and Horizon perform equally well and much better than JIT for constant offsets. Varying offsets have only minor impact on JET but significantly degrade the performance of Horizon.

As an important result in a two-class OBS node, we quantified the strong dependence of the high priority burst loss probability on QoS offsets and distribution of low priority bursts. This is the basis for designing aggregation strategies that assemble optical bursts at the edge of the optical network. Moreover, we showed that in principle the presented results also hold if the number of wavelengths is varied. Here, only the order of magnitude of losses changes.

Further work could extend the analysis for multiple classes and consider contention resolution such as buffering or deflection routing. Furthermore, a discussion on the behaviour of optical burst switching with various reservation mechanisms in a network scenario is necessary.

7 References

[1] Amstutz, S. R.: Burst switching - an introduction. IEEE Communications Magazine, November 1983, pp. 36-42.

[2] Amstutz, S. R.: Burst switching - an update. IEEE Communications Magazine, September 1989, pp. 50-57.

[3] Awduche, D.; Malcolm, J.; Agogbua, J.; O'Dell, M.; McManus, J.: Requirements for traffic engineering over MPLS. IETF, RFC 2702. September 1999.

[4] Awduche, D. et al., Multi-Protocol Lambda Switching: Combining MPLS Traffic Engineering Control With Optical Crossconnects. IETF, Draft draft-awduche-mpls-te-optical-02, July 2000. Work in progress.

[5] Batchelor, P.; Daino, B.; Heinzmann, P.; Hjelme, D.R.; Inkret, R.; Jäger, H.A.; Joindot, M.; Kuchar, A.; Le Coquil, E.; Leuthold, P.; de Marchis, G.; Matera, F.; Mikac, B.; Nolting, H.-P.; Späth, J.; Tillerot, F.; Van Caenegem, B.; Wauters, N.; Weinert, C.: Study on the implementation of optical transparent transport networks in the European environment - results of the research project COST 239. March 2000, pp. 15-32.

[6] Blake, S.; Black, D.; Calson, M.; Davies, E.; Wang, Z.; Weiss, W.: An Architecture for Differentiated Services. IETF, RFC 2475. December 1998.

[7] Braden, R.; Clark, D.; Shenker, S.: Integrated Services in the Internet Architecture: an Overview. IETF, RFC 1633. July 1994.

[8] Chaskar, H. M.; Verma, S.; Ravikanth, R.: A framework to support IP over WDM using optical burst switching. Proceedings of the Optical Networks Workshop, Richardson, Texas, 2000.

[9] Crovella, M.E. ; Bestavros, A.: Self-similarity in world wide web traffic: evidence and possible causes. IEEE/ACM Transactions on Networking. Vol. 5, No. 6, December 1997, pp. 835-846.

[10] Dolzer, K.; Gauger, C.; Späth, J.; Bodamer, S.: Evaluation of Reservation Mechanisms in Optical Burst Switching. AEÜ International Journal of Electronics and Communications. Vol. 55, No. 1, 2001.

[11] Dolzer, K.; Payer, W.: On aggregation strategies for multimedia traffic. Proceedings of the First Polish-German Teletraffic Symposium PGTS 2000, Dresden, September 2000.

[12] Ghani, N.: Integration strategies for IP over WDM. Proceedings of the Optical Networks Workshop, Richardson, Texas, 2000.

[13] Ghani N: Lambda-Labeling: A framework for IP-over-WDM using MPLS. Optical Networks Magazine, Vol. 1, No. 2, April 2000, pp. 45-58.

[14] Green, P. E.: Fiber optic networks. Englewood Cliffs: Prentice-Hall, 1993.

[15] O'Reilly, P.: Burst and fast packet switching: performance comparisons. Computer Networks and ISDN Systems, No. 13, 1987, pp. 21-32.

[16] Qiao, C.; Yoo, M.: Optical burst switching (OBS) - a new paradigm for an optical internet. Journal of High Speed Networks, No. 8, 1999, pp. 69-84.

[17] Qiao, C.; Yoo, M.: Choices, features and issues in optical burst switching. Optical Networks Magazine, Vol. 1, No. 2, April 2000, pp. 36-44.

[18] Qiao, C.: Labeled optical burst switching for IP-over-WDM integration. IEEE Communications Magazine, Vol. 38, No. 9, September 2000, pp. 104-114.

[19] Ramaswami, R.; Sivarajan, K.: Optical networks: a practical perspective. San Francisco: Morgan Kaufmann Publishers Inc., 1998.

[20] Ryan, J. P.: WDM: North American deployment trends. IEEE Communications Magazine, Vol. 36, No. 2, February 1998, pp. 40-44.

[21] Späth, J.: Dynamic routing and resource allocation in WDM transport networks. Computer Networks, Vol. 32, May 2000, pp. 519-538.

[22] Späth, J., Bodamer, S.: Routing of dynamic Poisson and non-Poisson traffic in WDM networks with limited wavelength conversion. Proceedings of the 24th European Conference on Optical Communication (ECOC '98). Regular and Invited Papers, Vol. 1, pp. 359-360, Lerko Prin S.A., Madrid, Spain, 20.09.1998 - 24.09.1998.

[23] Stern, T. E.; Bala, K.: Multiwavelength optical networks: a layered approach. Reading MA: Addison-Wesley, 1999.

[24] Turner, J. S.: Terabit burst switching. Journal of High Speed Networks, No. 8, 1999, pp. 3-16.

[25] Viswanathan, A., Feldman, N.; Wang, Z.; Callon, R.: Evolution of Multiprotocol Label Switching: IEEE Communications Magazine, Vol. 36, No. 5, May 1998, pp. 165-173.

[26] Wei, J. Y.; Pastor, J. L.; Ramamurthy, R. S.; Tsai, Y.: Just-in-time optical burst switching for multi-wavelength networks. Proceedings of 5th International Conference on Broadband Communications (BC’99), 1999, pp. 339-352.

[27] Yoo, M.; Qiao, C.: Supporting multiple classes of services in IP over WDM networks. Proceedings of IEEE GLOBECOM 99, December 1999, pp. 1023-1027.

[28] Yoo, M.; Qiao, C.; Dixit, S.: QoS performance in IP over WDM networks. IEEE Journal of Selected Areas in Communications, Special Issue on Protocols for the Next generation Optical Internet, Vol. 18, No. 10, October 2000, pp. 2062-2071.