Reservation Techniques in an OpMiGua Node

Andreas Kimsas, Steinar Bjornstad, Harald Overby and Norvald Stol

Norwegian University of Science and Technology (NTNU)
O. S. Bragstads plass 2B, 7491 Trondheim, Norway
e-mail: akimsas@item.ntnu.no

Abstract. An OpMiGua node integrates a packet switch for low priority traffic and a circuit switch for high priority traffic. Both traffic classes share the same input and output ports using time division multiplexing, but absolute priority is given to circuit switched packets. The circuit switched packets do not experience contention at output ports and are not subject to delay jitter; hence a guaranteed service class is created. Previous studies of OpMiGua nodes have used one of two reservation techniques to assure priority; either a time-window approach or a preemptive approach. This article introduces two new reservation techniques and investigates advantages and drawbacks associated with the four techniques. It is shown that each reservation technique is associated with specific loss mechanisms and methods to reduce their influence are proposed. Simulation results demonstrate that the choice of reservation technique is highly influenced by the relative share and length of high priority packets.

Keywords: OpMiGua, hybrid, time-window, preemption, packet loss

1 Introduction

Different hybrid Optical Burst/Circuit Switched (OBS/OCS) and Optical Packet/Circuit Switched (OPS/OCS) networks have been introduced over the last years. A survey of these networks is given in [1], where OpMiGua[1] [2] belongs to the "integrated" type of hybrid networks. Common for the integrated hybrid network approaches is the ability to share the bandwidth of a wavelength between packet switched and circuit switched traffic. The time-scale for resource sharing is on a packet-by-packet basis and is different from hybrid schemes where resources are reserved for timescales equalling circuit switched connection times. This capability is used to increase the resource utilisation as compared to pure OCS networks.

The OpMiGua concept was developed to create a network architecture that is capable of achieving high resource utilisation *and* capable of providing a service class with guaranteed delivery. Guaranteed service class traffic (GST) is created by employing optical circuit switched paths to avoid contention in transit nodes and to eliminate delay jitter. To increase the bandwidth utilisation, statistically multiplexed (SM) traffic is injected in the voids between consecutive GST packets. Reservation techniques are used to assure that the insertion mechanism does not interfere with the GST packets. It will be shown that the reservation techniques used to assure priority

[1] Optical Migration Capable Network with Service Guarantees

for GST packets are associated with loss mechanisms that influence the performance of the SM traffic. With the final goal of reducing the blocking probability of SM traffic, this article investigates the loss mechanisms and performance of different reservation techniques.

Earlier articles on OpMiGua suggested using reservation techniques based on a time-window approach [2] or based on a preemptive approach [3]. In addition to the two previous techniques, two new reservation techniques are introduced; an improved time-window approach and a combined time-window and preemptive approach. Comparison of the four reservation techniques will provide insight into the loss mechanisms involved and the performance of the two new reservation techniques can then be compared to the existing ones. With a focus on reservation techniques buffering for contention resolution is not considered.

Each technique is evaluated via simulations of an OpMiGua node, whose operation is briefly explained in section 2. Section 3 discusses the logical behaviour of the four reservation techniques and identifies the loss mechanisms involved in each method. The simulation scenario and results are treated in section 4 and the main findings are then summarised in the conclusion.

2 The OpMiGua Node

An OpMiGua core node can be decomposed into four main parts. A packet separator (PS) detects if the incoming packets belong to the GST class or to the SM class by analysing its optical label. Possible labelling methods include use of orthogonal polarisations [4] or subcarrier modulation [5]. Regardless of the separation technique in use, the packet separator will forward GST packets to an optical cross connect (OXC) and SM packets to an optical packet switch (OPS). Performance of a blocking OPS module was studied in [3], but in this article both OXC and OPS modules are assumed strictly non-blocking.

The wavelength routed GST packets are directed to its output port according to the current configuration of the OXC, while the SM packets can use any idle output wavelength which is part of the correct output fibre. The insertion of SM packets is performed by an optical packet switch inside the OpMiGua node. Two questions then arise; how can the arrival of GST packets be detected and signalled, and how can absolute priority be ensured for GST traffic? The answer to the first question is given below, while the second question is treated in section 3.

The packet combiner (PC) is responsible for signalling the presence of incoming GST packets to the control unit of the packet switch and re-inserts the optical label if necessary. The physical detection of GST packets is enabled by an optical splitter, which is linked to the control unit. After passing the optical splitter, the GST packets enter a fibre delay line (FDL) whose purpose is to create a time-window for the control unit. The size of the time-window is determined by the length of the FDL and its actual length depends on the reservation techniques in use. Having knowledge of future GST arrivals, the control unit can make the necessary decisions to avoid simultaneous transmission of GST and SM packets on the same shared output port.

3 Logical behaviour of reservation techniques

The time-window technique [6] [7] is used in OpMiGua to predict and avoid simultaneous transmission of GST and SM packets on the same wavelength and is a *proactive* scheme. The preemptive technique [8] [9] avoids simultaneous transmission by interrupting SM transmission after a conflict is detected and is a *reactive* approach. As will be explained in this section, each method has their advantages and disadvantages with respect to network layer efficiency and implementation complexity.

3.1 Time Window Techniques

Upon arrival of an SM packet at the OPS switch, the control logic is aware of all GST arrivals within the time $t_{arrival}+\Delta$, where Δ is the size of the time-window and equals the transmission time for the longest SM packet. Provided that SM packets are scheduled to an output port whose time-window is empty there is no risk of contention with GST packets. Figures 1 and 2 illustrate the logical behaviour of the two time-window techniques considered in this article; the simple time window (STW) scheme and the length aware time window (LATW) scheme. For simplicity, only one output wavelength is shown. Also, contention among SM packets is supposed to be handled by the OPS module which explains the absence of SM packets that overlap in time.

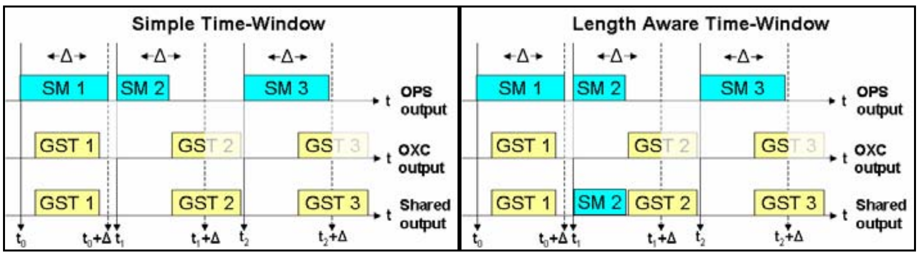


Fig. 1 left (*STW*) and Fig. 2 right (*LATW*): The blocking probability is higher for the STW technique, but its implementation is simpler as it does not take SM packet length into account.

The STW technique does not consider the length of SM or GST packets when scheduling SM packets. The control-logic will schedule SM packets on the output port as long as GST packets are not detected within the time-window. Referring to the figure, at time t_0 a scheduling decision has to be made for SM-1. The control logic has knowledge of the arriving GST-1 packet which should be given priority. SM-1 is then inserted into a buffer, scheduled on an alternative available wavelength or discarded. The same occurs for SM-2 at time t_2 and SM-3 at time t_3. However, as exemplified by SM-2 in the figure, there are inefficiencies related to this reservation technique. Although SM-2 is sufficiently short to be scheduled without disturbing GST-2 it was not feasible since the control-logic does not take packet length into account. This inefficiency has been named reservation induced blocking (RIB) [2] and adds to blocking caused by contention of simultaneously arriving SM packets. Thus, at the expense of RIB absolute priority to GST packets is achieved.

The LATW technique is suggested to reduce the RIB. For each output wavelength, the time to arrival of the GST packets is continuously updated so that the control-logic can compute whether there is sufficient time to schedule the SM packet on one of the outputs. Either it can schedule the packet on the first wavelength that has sufficient space, or it goes through all wavelengths and schedules the packet on the wavelength which creates the smallest gap between the SM and GST packet. On average, the first option requires less computation, but the second option makes better use of each wavelength resource. Our simulations use the minimum gap method, but whichever case being chosen the RIB is eliminated because SM packets are only rejected when there is contention at the output port. Also, the blocking probability for any packet length using the LATW scheme will be inferior or equal to the STW scheme. This scheme can therefore be seen as the optimal choice with respect to performance, but requires access to the SM packet length and involves more computation. Additionally, in contrast to the STW technique blocking probability increases for longer SM packet lengths which introduce unfairness.

3.2 Preemptive and Combined Techniques

The preemptive (PRE) technique depicted in figure 3 does not rely on a time-window, but will always attempt to transmit the SM packet if the shared output wavelength is currently available. At time t_0, the control logic has no knowledge about a possible conflict between SM-1 and GST-1, so SM-1 is scheduled on the common output wavelength. Once the GST packet is sensed on the input, transmission of SM-1 is immediately terminated and SM-1 is said to be preempted. Fragments of preempted packets will cause an increased load at the subsequent nodes, unless precautions are taken to detect and remove the fragments.

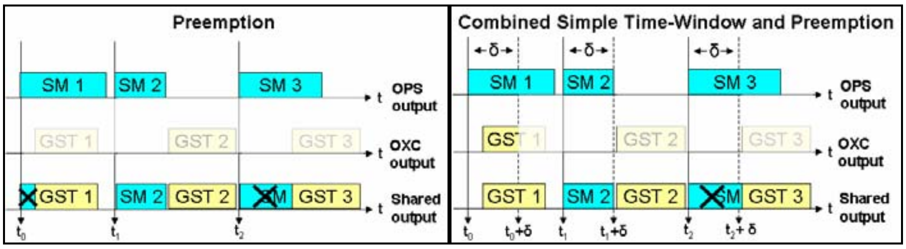


Fig. 3 left (*PRE)* and Fig. 4 (*COMB*): The preemptive technique allows SM packets to be scheduled on any free output port at the risk of being preempted. A cross indicates preempted packets. With the combined scheme SM packets which are longer than the time window δ risk preemption, while shorter packets may be blocked by incoming GST packets.

In contrast to the STW technique, SM-2 is scheduled at time t_1 and is successfully inserted between GST-1 and GST-2. Although not suffering from RIB, the performance of the preemptive scheme is affected by a different loss mechanism. If several outputs wavelengths are free, the control logic must schedule the packet on one of these outputs without knowledge about future GST arrivals. Hence, it is possible that the selected output port results in preemption, even though it could have been successfully transmitted on a different output port. This inherent inefficiency for the preemptive scheme can bee seen as a loss of multiplexing gain and is termed preemption induced blocking (PIB).

When considering the length of SM-2 it should be clear that the probability of a successful transmission decreases when the packet-length is increased. Again this is exemplified with SM-3, which would not have been preempted if it was of the same length as SM-2. Hence, the preemptive scheme introduces unfairness with respect to packet length, where shorter packets are favoured.

The combined (COMB) scheme use elements from the STW and PRE techniques. With this method we try to find a compromise between the amount of RIB and PIB, as well as reducing the number of preempted packets. The scheme use a time window of size δ, with $0 < \delta < \Delta$. It should be noted that the combined scheme equals the STW scheme for $\delta = \Delta$ and the preemptive scheme for $\delta = 0$. As shown in figure 4, the time window is sufficiently large to predict collision between SM-1 and GST-1. When compared to the STW scheme, the successful scheduling of SM-2 illustrates a reduced RIB. Since Δ is the minimum length required to assure collision avoidance, the shorter time window implies that all scheduled packets with length between δ and Δ may interfere with incoming GST packets. If interference is detected at the output, then absolute priority to GST packets is obtained by preemption, illustrated by SM-3 in the figure.

4 Simulation Scenarios and Results

4.1 Simulation Scenarios

The performance of each scheme is assessed via simulation of an all-optical OpMiGua node, with its main parameters listed in table I. Strictly non-blocking OPS and OXC modules is assumed to get the full benefit of statistical multiplexing for SM packets and eliminate blocking probability for GST packets. Both SM and GST packets are generated according to a Poisson arrival process. All simulation scenarios in this article assume that the input traffic is buffered in ingress nodes, i.e. packets that overlap in time will be transmitted back-to-back. The simulation points were obtained running ten independent replications to establish 95% confidence intervals and the simulation length was adjusted to produce confidence intervals which are smaller than 10% of the average value. To improve the readability of the figures the confidence intervals are not represented, but their values are included in the discussion whenever necessary.

According to the connection-oriented nature of GST packets, all packets generated at an input port are directed to an output port which is not used by other GST sources. The destination fibre for SM packets is uniformly distributed among the four fibre outputs. While SM packets are transmitted one-by-one, earlier studies [2] have shown that aggregating GST packets into bursts at the ingress nodes have positive effects for the SM packet loss rate (PLR). Consequently, most of the simulations set the burst length to $100 \cdot SM_{MEAN}$, which is thought to be a good compromise between performance and aggregation delay. The aggregation process itself was not simulated, as a thorough study of different GST burst aggregation schemes will be investigated in a separate paper.

Table I: Parameters for Simulation Study. Brackets indicate a list of possible values.

Component / Quantity	Parameter	Parameter value	Unit
Bitrate per wavelength	Rate	10	[Gbps]
Fibre inputs / outputs	F	4	
Wavelengths per fibre	W	32	
Normalised load	A	$0.6 \leq A \leq 0.9$	
GST share of total load	S	$0 \leq S \leq 100$	[%]
SM packet length	SM_{LENGTH}	Uniform 40-1500	[bytes]
SM average packet length	SM_{MEAN}	770	[bytes]
SM maximum packet length	$SM_{MAX\ LENGTH}$	1500	[bytes]
GST burst length	GST_{LENGTH}	$[1,10,100,1000] \cdot SM_{MEAN}$	[bytes]
Time-window	Δ	$SM_{MAX\ LENGTH} \cdot 8$ / Rate	[s]

4.2 Influence of GST share on the different reservation techniques

This section aims at quantifying the effect of the different loss mechanisms related to the SM class, particularly in presence of varying GST shares. Being statistically multiplexed, the SM packets will experience contention at the output ports whenever all possible output wavelengths are occupied by SM or GST packets. Throughout the paper this loss mechanism is termed contention induced blocking (CIB) and is present for all reservation techniques studied here. The figures 5-8 show the SM PLR for different loads and for different shares of GST traffic.

Time Window Techniques

When considering the STW technique, the total loss is the result of the CIB and the RIB. The RIB equals zero in the absence of GST traffic, which implies that the system is purely packet switched, the curve for GST share S=0% in figure 5 is then used as reference for the curves with non-zero GST share.

Using experience for purely packet switched systems, the introduction of two service classes normally introduces inefficiencies, hence an increase in PLR could have been expected [8]. In figures 5 and 6 SM PLR clearly benefits from high GST shares for loads inferior to 0.8, thus an OpMiGua node with two service classes performs better than a single service class OPS node. So, for low and moderate loads, the negative effect of RIB is more than compensated for by the reduced CIB. When the load exceeds 0.8 this is no longer true. Take the points at A=0.9 as examples; the upper value for the S=0% curve equals 0.062 while the lower value for S=90% equals 0.081. Here the output ports are so congested that blocking due to RIB overshadows the reduced CIB. A confirmation of this analysis is found in figure 7, where the RIB is reduced by employing the LATW technique. The curves around A=0.9 almost overlap within the confidence intervals.

The LATW technique was proposed to eliminate the effect of RIB. As expected this leads to better performance regardless of the load. When comparing the curves of equal GST share in figures 5 and 6 the difference is small, but still present. For

S=30% at load 0.65 the PLR is improved with 4.7 %, but for S=70% the improvement is 24.4 %. Similarly, at load 0.9 the gain increase from 5.1% to 18.2%. The trend is clear: higher GST share gives higher improvement. This is in accordance with the fact that with the STW technique the RIB will be more and more prevalent as the amount of GST traffic increases, so reducing the RIB will favour the simulation scenarios with high GST share.

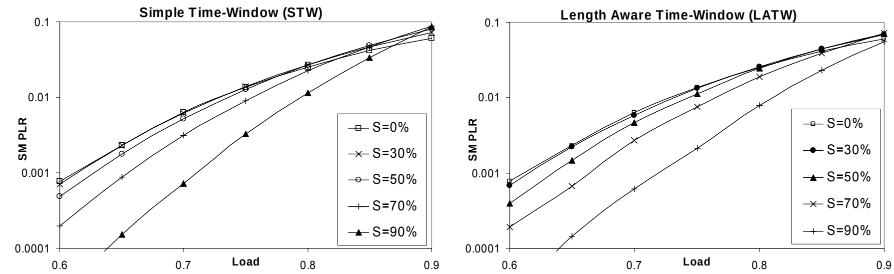


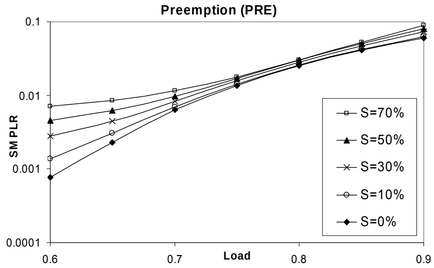
Fig. 5 left (*STW*) and Fig. 6. right (*LATW*). $GST_{LENGTH} = 100 \cdot SM_{MEAN}$. For loads A<0.85 the SM PLR benefits from high GST shares. The SM PLR for the LATW technique is improved by up to 20% for the curves represented here.

The effect of the time window approaches can then be summarised. First, the SM traffic in an OpMiGua node with moderately aggregated GST traffic performs better than a purely packet switched node with a single service class. Second, since the GST traffic bypass the OPS switch, the OPS switch will experience a reduced *internal* load as compared with a purely packet switched node. Hence, the SM class in OpMiGua is less influenced by internal blocking when compared to low priority traffic in a purely packet switched node. This property can be exploited by using a cost-effective OPS module with internal blocking, rather than the more expensive and technologically immature non-blocking alternatives [3]. Finally, a relatively low PLR is obtained for the SM class, even without the use of buffering for contention resolution. This is especially true for high GST shares. Hence, it could be beneficial to let part of the SM traffic be sent as GST packets if the traffic demand on an optical lightpath is low. However, the performance and implementation of this last issue is outside the scope of the article, and is left for future studies.

Preemptive and Combined Techniques

The SM PLR represented in figure 7 shows a behaviour that is opposite to the behaviour using time-window techniques; the performance deteriorates with increased GST share. Again, the total PLR is the result of two effects, CIB and PIB. The analysis for the time-window techniques showed that the CIB decreases as more GST traffic is inserted into the system. Since the simulation scenarios only differ with respect to the reservation technique, the same should apply here. Even more, with a higher packet loss the average occupancy probability of the output port will decrease, so the CIB should be even lower than for the time-window techniques. Consequently, the increased PLR is caused by the PIB.

At low loads, the SM PLR for curves with high GST shares is very different from the case without GST traffic. One may conclude that for high GST shares the loss mechanism is primarily influenced by the PIB, while for low GST shares the primary loss mechanism is the CIB. As the total load increases towards A=0.8 the curves almost overlap regardless of GST share. Therefore, in this regime the CIB is the dominant loss mechanism. At even higher loads the curves split, with the highest loss for the highest GST shares. Both CIB and PIB increase, but the PIB increases faster than the CIB.



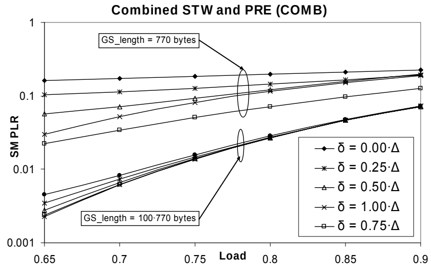


Fig. 7 left (*PRE*) and Fig. 8. right (*COMB*): The preemptive technique for different GST traffic shares (S) and the combined scheme with different time-windows δ and S=30%.

Some general observations can be made concerning the performance of the preemptive technique. For all loads it is outperformed by the LATW technique, and for loads inferior to 0.9 it also is less performing than the STW technique. However, the difference is not considerable for GST shares inferior to 30%. At load 0.7 and S=30%, the PLR for the preemptive technique is $8{,}2\cdot10^{-3}$ compared to $6{,}2\cdot10^{-3}$ for the STW technique.

The combined technique is evaluated by varying δ in the interval [0, Δ]. For δ=0 the reservation technique is entirely preemptive and for δ=Δ the scheme corresponds to the STW technique. When the GST packets are aggregated into burst of 100, then all curves for the combined scheme in figure 8 are confined within the limits defined by the curves for δ=0 and δ=Δ. If packet fragments are not considered, the choice of reservation technique is reduced to finding a compromise between the SM packet loss rate and added GST delay at the intermediary nodes. Experience from the preemptive technique and the STW technique indicates that the difference between the curves will be reduced for lower GST shares and increased for higher GST shares. At higher loads the difference between the reservation techniques will be small, regardless of the GST share and regardless of the scheme being used.

When considering the case without GST packet aggregation, cf. the five upper curves in figure 8, the performance of the combined scheme with δ = 0.75·Δ outperforms the STW technique. It is not the case for other values of δ. So, by allowing a well chosen ratio of packets to risk preemption, the performance is improved.

The analysis illustrates the complexity of the packet loss mechanisms in an OpMiGua node. Hence, a large parameter space is required to get a good understanding of the system. Furthermore, the two set of simulations in figure 8 also illustrate the sensitivity on GST burst length, an issue which is studied in next section.

4.3 Effect of Burst Length

If the STW technique is used, an arriving GST burst will contribute the RIB during the time interval defined by the time of entry into the FDL and the arrival time at the packet combiner. The length of the burst is of no importance when considering RIB, better performance can then be expected if the average number of burst arrivals is reduced. For a given GST share, this may be achieved by increasing the burst length. Indeed, significant improvement in the SM PLR is observed for the STW technique in figure 9, especially when moving from no aggregation to an aggregation length of 10 GST packets. The effect drastically decreases after an aggregation length of 100 GST packets, almost nothing is gained when shifting to an aggregation length of 1000 packets. In other words, the RIB is already much smaller than the CIB for a GST_{LENGTH} equal to 100 packets. Unless buffers for contention resolution are used, extending the burst length will only imply added delay.

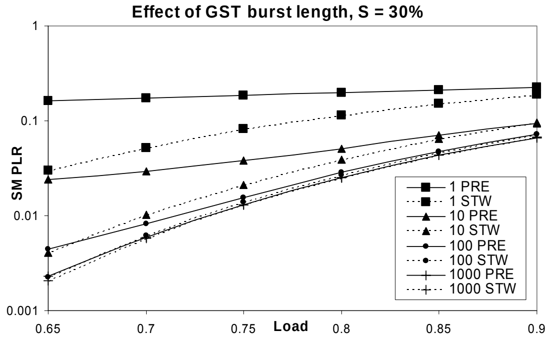


Figure 9: Effect of burst length for the preemptive technique (*PRE*) and the simple time-window technique (*STW*) . S=30%.

The same trend applies for the preemptive technique, but the difference between the simulation scenarios is larger than for the STW technique. As explained in the section on the preemptive approach, PIB is the dominating loss mechanism at low and moderate loads. Reducing the number of GST burst arrivals reduce the influence of PIB. For the 100 STW scenario the RIB is not dominant compared to the CIB, but for the preemptive technique this is not the case. A substantial improvement can still be observed when increasing the aggregation length to 1000 GST packets, illustrating that the preemptive technique needs more GST aggregation than the STW technique to achieve the same SM PLR.

5 Conclusion

In this paper we have identified and proposed reservation mechanisms for the OpMiGua hybrid optical node. The novel length aware time window and the combined preemption/time-window schemes have been thoroughly investigated. Circuit switched guaranteed traffic (GST) has no packet loss. Blocking probability of low priority traffic is used as performance parameter.

We found that regardless of reservation technique, the blocking probability is heavily influenced by the share of GST traffic, and improves as the GST burst length increases. Hence, a compromise between GST aggregation delay and packet loss must be found when designing an OpMiGua network.

For traffic loads less than 0.8, the STW and LATW techniques outperform the preemptive and combined techniques. However, for GST shares inferior to 30% the difference between the techniques is small, e.g. less than 10% for a GST share of 10% and a load of 0.7.

For traffic loads beyond 0.8 the difference in performance decreases, for all traffic loads beyond 0.8 the difference in performance decreases, for all schemes and for all GST traffic shares.

Furthermore, a higher performance than for an OPS with a single service class is found using both the STW and the LATW techniques, given a load below 0.8 and a GST shares above 30%. This illustrates the potential of the OpMiGua concept and motivates a future study of scheduling approaches where SM packets are transmitted as GST traffic on circuits with low GST share.

Acknowledgment: Telenor, the NRC and COST action 291 are recognised for funding research activities in the field of optical networking.

References

1. Gauger C.M., Kühn P.J., Breusegem E., Pickavet M., Demeester P.: Hybrid Optical Network Architectures: Bringing Packets and Circuits Together. IEEE Communications Magazine (August 2006), vol. 44, pp. 36-42.
2. Bjornstad S., Hjelme D.R., Stol N.: A Packet-Switched Hybrid Optical Network with Service Guarantees, IEEE JSAC, Supplement on Optical Communications & Networking, (August 2006), vol. 24, pp. 97-107.
3. Kimsas A., Bjornstad S., Overby H., Stol N.,: Protection Using Redundancy in a Hybrid Circuit/Packet Node Design. Proc. ECOC (2006), We3.P.143.
4. Tuft V.L., Bjornstad S., Hjelme D.R.: Time Interleaved Polarization Multiplexing for Polarization Labelling, Proceedings of the 7th International Conference on Transparent Optical Networks (July 2005), vol. 1, pp. 47-51.
5. Breusegem E. et al.: A B Broad view on Overspill Routing in Optical Networks: a Real Synthesis of Packet and Circuit Switching?", Journal of Optical Systems and Networking (2005), vol. 1, no. 1, pp. 51-64.
6. Hu G., Gauger C.M., Junghans S.: Performance of MAC Layer and Fairness Protocol for the Dual Bus Optical Ring Network (DBORN), Conference on Optical Network Design and Modeling (February 2005), pp. 467-476.
7. Qiao C., Yoo M., Dixit S.: OBS for Service Differentiation in the Next-Generation Optical Network. IEEE Communications Magazine (February 2001), vol. 39, no. 2, pp. 98–104.
8. Overby H., Nord M., Stol N.: Evaluation of QoS Differentiation Mechanisms in Asynchrounous Bufferless Optical Packet Switched Networks. IEEE Communications Magazine (August 2006), vol. 44, no. 8, pp. 52-57.
9. Phuritatkul, J., Yusheng J.: Analysis of Bandwidth Allocation with Wavelength Preemption Scheme for Supporting Relative Service Guarantees in OBS networks. Broadband Networks (October 2005), vol. 2, pp. 1300 – 1309.