of 10 (achieved by decreasing t_{mem} and t_{ALU} by an order of magnitude) also does not lead to an important improvement: networks of 35-65 nodes can be supported in this case. Computationally simpler SP-FF could be used to increase the scalability of WR-OBS (50-115 nodes with current processors), but due to its poor blocking probability [5] it may be sub-optimal for implementation. These results show that centralized WR-OBS could be applied to medium-size networks, as metropolitan-area networks.

3. Dynamic WR-OBS versus static WRON

To investigate under which conditions WR-OBS can save wavelength resources with respect to WRONs, a WR-OBS network was modelled as core architecture, applied, by way of an example to the EUROCORE [4] and NSFNET [2] topologies. Both networks were equipped with the same number of wavelengths required in the static WRON [2] (4 and 13 wavelengths in the links with the highest wavelength requirements, respectively) and AUR-E was implemented in the central node because of its lowest blocking probability. The maximum offered load at which the request blocking probability is lower than 10^{-4} was then determined through simulation. Next, the network capacity was gradually decreased removing the least used wavelength in the network from the corresponding links and the maximum offered load was calculated again. The input traffic at each buffer was assumed as an ON-OFF process where the ON and OFF periods are Pareto distributed with parameter $\alpha = 1.5$ to model self-similar traffic. The ON period was set so the efficiency requirement in a dynamic network mentioned in the previous section is met. The results, in Fig. 3, show that in both networks wavelength savings occur at low loads. Higher savings without increasing the network capacity could be achieved implementing a more computationally complex central node: it has been shown that using information about the already scheduled bursts [7] or a Earliest Deadline First scheduling scheme with retrials instead of FIFO [3] decrease the blocking probability significantly. However this would lead to a reduction in achievable scalability. For instance, if retrials are allowed and every request is processed twice, scalability reduces to 15-30.

4. Summary

Analytic equations were derived, for the first time, to quantify the scalability of centralized wavelength-routed dynamic network architecture. Results show that the maximum number of nodes supported by such networks with today's processors is of the order 20-115 depending on dynamic RWA used, network topology and diameter - which makes WR-OBS suitable for medium-size networks. In addition, new results show that WR-OBS uses fewer wavelengths than static WRONs at low traffic loads. To achieve wavelength savings at higher loads -without increasing the network capacity, extra functionalities similar to those proposed in [3,7] must be implemented in the central node. However, requiring the central node to perform more tasks prior to lightpath allocation leads to a reduction in scalability, highlighting a clear trade-off between scalability and resource savings. These results define the operating limits for a centralized wavelength-routed dynamic network and are applicable to the design of dynamic and static wavelength-routed network architectures.

5. References

[1] M. Düser, P. Bayvel, *J. of Lightwave Techn., Vol. 20*, 574-585 (2002).
[2] S. Baroni, P. Bayvel, *IEEE J. of Lightwave Techn., Vol. 15*, 242-251 (1997).
[3] E. Kozlovski *et al., Proc. IEEE LEOS 2001 Annual Meeting*, Nov. 2001, paper TuD2.
[4] M. Garnot et al., *Proc. ICC'97*, 422-426.
[5] A. Mokhtar, M. Azizoglu, *IEEE/ACM Trans. on Networking, Vol. 6*, 197-206 (1998).
[6] "The microarchitecture of the Pentium 4 processor," (*Intel, 2001*). http://developer.intel.com/technology/itj/q12001/pdf/art 2.pdf.
[7] K. Dolzer *et al., AEU Int. J. Electron. Commun. 55* No. 1, 1-8, (2001).

TuJ2 **2:15 PM**

Intermediate Node Initiated (INI) Signalling: A Hybrid Resevation Technique for Optical Burst-Switched Networks

R. Karanam, V. Vokkarane, J. Jue, *The University of Texas at Dallas, Richardson, TX, Email: vinod@utdallas.edu.*

We introduce a new reservation scheme, Intermediate Node Initiation, which provides the functionality of both TAW and JET in the same network. INI scheme provides a trade-off between end-to-end delay and burst loss, enabling support for different classes of traffic.

1 Introduction

Optical burst switching (OBS) is a new paradigm proposed to efficiently support the ever-growing broadband traffic of the Internet directly over all-optical WDM networks [1]. A significant component of OBS networks are the signaling and reservation scheme [2]. Two prominent reservation schemes for a bufferless OBS network are Tell-and-Wait and Just-Enough-Time. In both of these schemes, a burst header packet (BHP) is sent ahead of the data burst in order to configure the switches along the burst's route.

In Tell-and-Wait (Fig. 1(a)), the BHP is sent along the burst's route to collect channel availability information at every node along the path. At the destination, a channel assignment algorithm is executed, and the reservation period on each link is determined based on the earliest available channel times of all the intermediate nodes. A *reply* packet is sent in the reverse direction, which then reserves the channel for the appropriate duration at each intermediate node. At any node along the path, if the required channel is already occupied, a *fail* packet is sent to the destination to release the previously reserved resources. If the reply packet reaches the source successfully, then the burst is sent at the scheduled time. TAW in OBS differs from wavelength-routed WDM networks in that, resources are reserved only for the duration of the burst, and no explicit disconnect message needs to be sent.

In Just-Enough-Time (Fig. 1(b)), the source node first sends a BHP towards the destination node. The BHP is processed at each subsequent node, resulting in an appropriate channel selection, and a reservation on that channel for the duration of the burst. If the channel reservation is successful, the switch will be configured immediately prior to the burst's arrival. Meanwhile, the burst waits at the source in the electronic domain. After an offset time, T, whose value is calculated based on the number of hops from source to destination and the switching time of a node, the burst is sent optically on the chosen wavelength [1]. If the reservation is unsuccessful at any node, then the burst will be dropped.

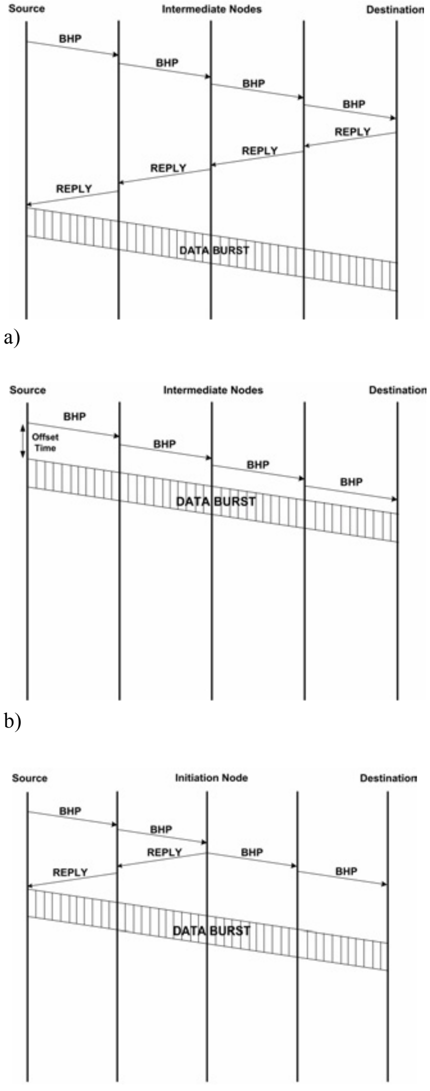
The disadvantage of TAW is the round-trip setup time, i.e., the time taken to set up the channel; however in TAW, the loss probability is very low due to the channel reservation acknowledgment. Therefore, TAW is suitable for delay insensitive traffic. In JET, the loss probability is high, but the end-to-end delay is less than TAW, since there is no need for an end-to-end acknowledgment. Thus, JET is suitable for loss insensitive traffic. Neither signaling schemes offer flexibility with respect to both delay and loss tolerance values.

In this paper, we propose a new reservation scheme, *Intermediate Node Initiated* (INI) signaling, which takes into account the advantages of both TAW and JET. The reservation request is initiated at an intermediate node, called the *initiating node* (IN). In the first part of the path, i.e., from source to the initiating node, the INI reservation scheme works with an acknowledgment for the BHP similar to TAW, and in the later part of the path, from the initiating node to destination, the INI reservation scheme works without an acknowledgment, similar to JET.

2 Intermediate Node Initiated (INI) Signaling

In the INI reservation scheme, a BHP is sent to the destination, and a node between source and destination is selected as the initiating node (Fig. 1(c)). The BHP collects the channel availability information at every node along the path until it reaches the initiating node. At the initiating node, a channel assignment algorithm is executed to determine the time duration that the channels will need to be reserved at each intermediate hop between the source and initiating node. A *reply* packet is then sent to the source node that reserves channels along the path from the initiating node to the source for the appropriate duration. If a channel is busy, a *fail* packet is sent back to the initiating node to release previously reserved resources. If the reply packet reaches the source successfully, then the burst is sent at the scheduled time. The BHP sent from the initiating node towards the destination, reserves channels for the burst in a similar manner as JET. If the BHP fails to reserve the channel at any node between initiating node and destination node, the burst is dropped at that node.

In INI, there is an acknowledgment from the initiating node to the source node, thereby decreasing the probability of loss compared to JET. INI also decreases the end-to-end delay as compared to TAW, since the burst waits at the source for a duration less than the round-trip propagation delay between the source to destination as in TAW. In INI, if the initiating node is the source node, the reservation scheme is identical to JET, and if the initiating node is the destination node, the reservation scheme is identical to TAW.



a)

b)

c)

Fig. 1. (a) Tell-and-Wait (TAW), (b) Just-Enough-Time (JET), (c) Intermediate-Node-Initiated (INI).

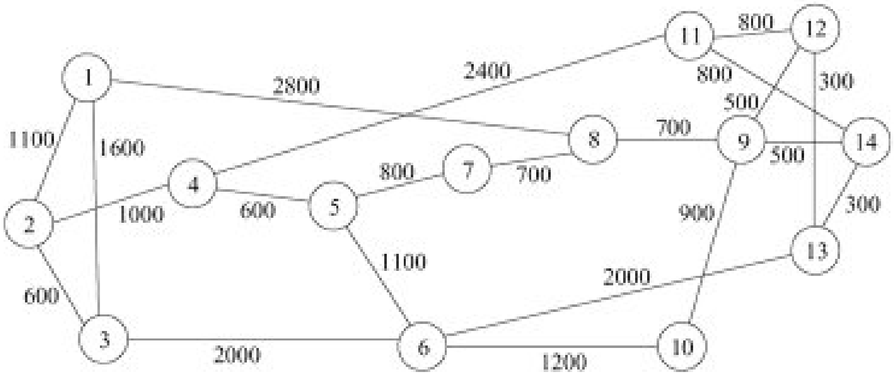


Fig. 2. NSFNET with 14 nodes.

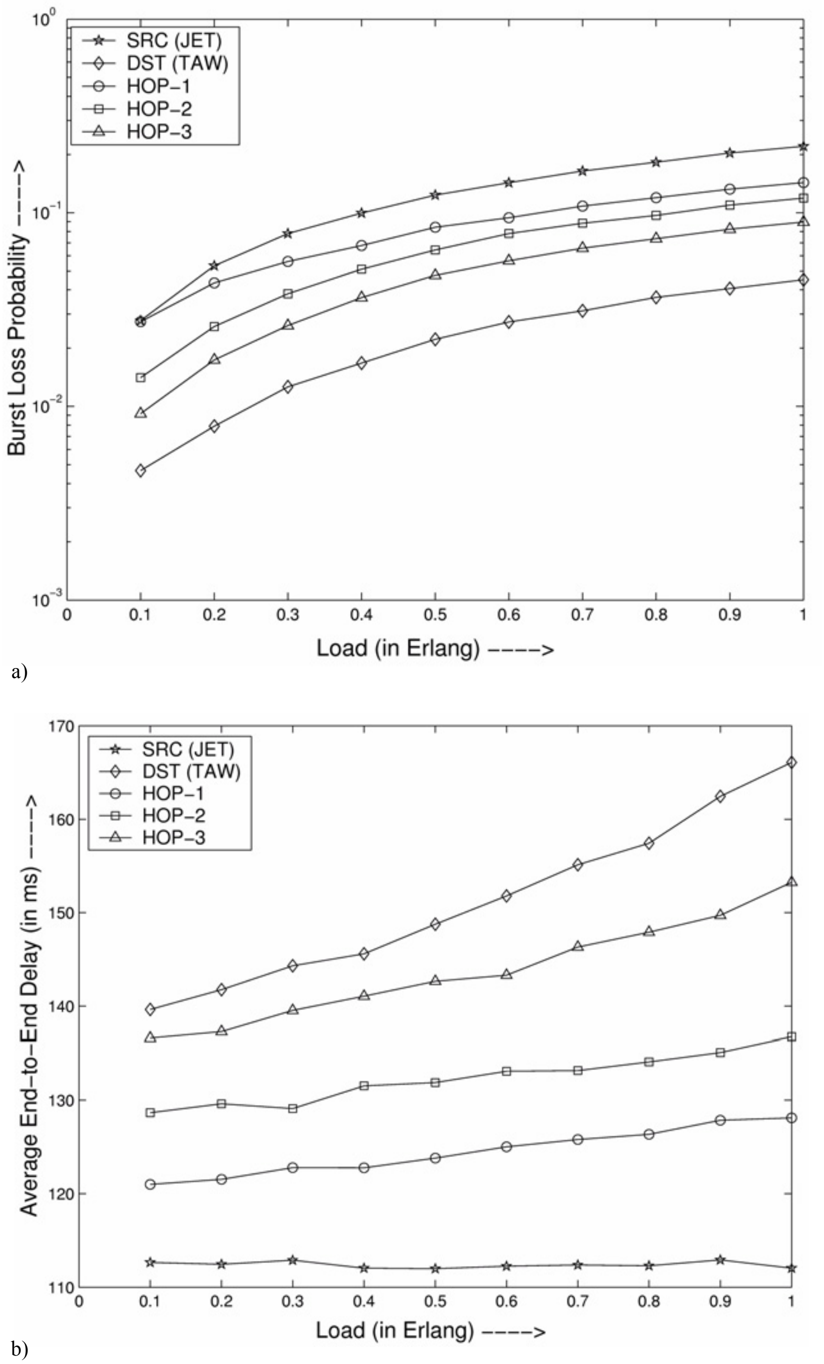


Fig. 3 (a) Burst loss probability versus load, and (b) Average end-to-end delay versus load, when the initiating nodes are source, first hop, second hop, third hop and destination.

Table 1 gives the summary of the three reservation schemes in terms of burst loss probability and average end-to-end delay. We can observe that the loss and delay for the new scheme is in-between the currently proposed schemes.

Table 1: Reservation Schemes for OBS

Reservation Scheme	Burst Loss Probability	Average end-to-end Delay
TAW	Low	High
JET	High	Low
INI	Medium	Medium

3 Simulation Results

In order to evaluate the performance of the INI scheme, a simulation model is developed. Burst arrivals to the network are Poisson with rate λ. Burst length is exponentially distributed with average burst length of $1/\mu = 100$ ms. Transmission rate is 10 Gbps. Packet length is 1250 bytes. Switching time is 10μs. There is no buffering or wavelength conversion at nodes. Retransmission of the lost bursts is not considered. Figure 2 shows the 14-node NSFNET on which the simulation is implemented. The distances shown are in km.

Figure 3(a) and 3(b) plot the burst loss probability and average end-to-end delay versus load when the initiating nodes are taken as source (SRC), first-hop (Hop-1), second-hop (Hop-2), third-hop (Hop-3), and destination (DST) respectively. In Figs. 3(a) and 3(b), only paths that are more than or equal to three hop count are considered to show the effect of INI scheme. We observe that the loss probability decreases as the initiating node moves away from the source. If the initiating node is chosen closer to the source, a greater part of the path is un-acknowledged, which leads to a higher loss probability. On the other hand, if the initiating node is chosen closer to the destination, a greater part of the path is acknowledged, which leads to a lower loss probability. We also observe that the delay increases proportionally to the increase in distance between the initiating node and the source, since the path from source to the initiating node is acknowledged, and hence incurs a higher round-trip delay. Also, the values of loss and delay when the initiating node is at the source and the destination are consistent with JET and TAW respectively.

The INI scheme can be extended to provide QoS at the optical layer. It is possible to implement multiple signaling schemes in the same network to provide differentiated services, in order to support both loss and delay sensitive traffic, i.e., we can use TAW for loss sensitive traffic, and JET for delay sensitive traffic. Using INI, we can satisfy both these constraints by carefully selecting the initiating node. Suppose we have to support three classes of traffic, say P0, P1, and P2, with P0 being delay sensitive, P1 being both delay and loss sensitive, and P2 being loss sensitive, then we can use the source node as the initiating node for P0, the center node as the initiating node for P1, and the destination node as the initiating node for P2, thus providing differentiated services in the same OBS network.

Figure 4(a) and 4(b) plot the burst loss probability and average end-to-end delay versus load for the three priority bursts. We observe that P2 suffers the least loss, while P0 incurs the least delay, and P1 is in-between the values of P0 and P2. For comparable values of offset-time, we found that INI out-performs the traditional offset-based QoS scheme [3]. Since, in the offset-based scheme the source has to estimate the additional-offset to provide differentiated services, while in INI, the initiating node has the channel availability information of all nodes between itself and the source. Also, the data burst does not enter the network until resources have been reserved between the source node and the initiating node.

4 Conclusion and Future Work

In this paper, we introduced the intermediate node initiated signaling scheme for an OBS network. The INI reservation scheme provides flexibility during channel reservation based on the type of data to be transmitted. The packet loss probability of INI is less than JET and the end-to-end delay is less than TAW. Hence, the proposed hybrid scheme is a flexible solution suitable for handling the varying traffic demands of the next generation optical network. An area of future work is to study the performance of the INI scheme with wavelength conversion and deflection to improve channel utilization. Also, the performance of INI can be improved by implementing void filling, i.e., utilizing channel gaps between existing reservations, for reservation.

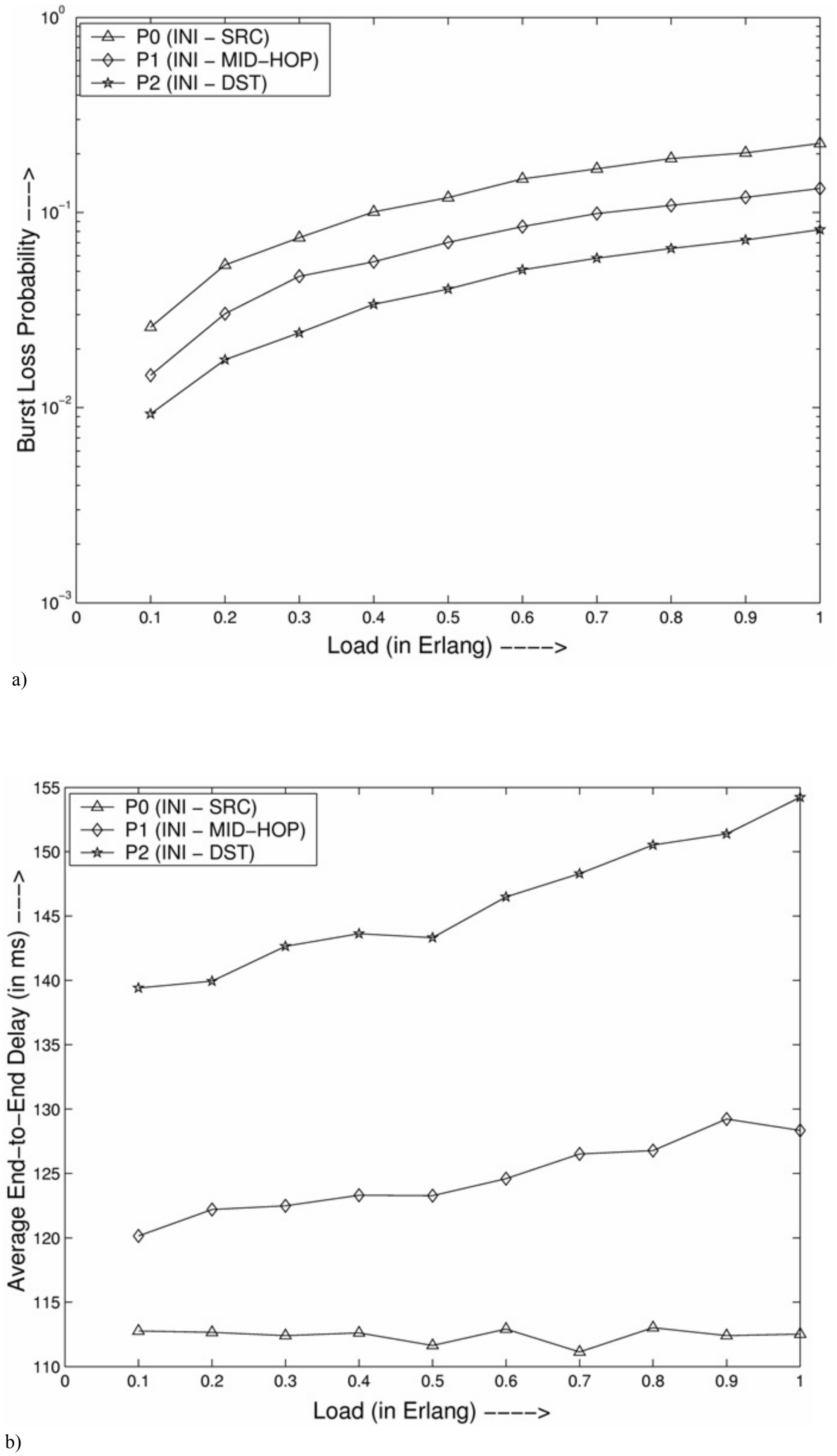


Fig. 4 (a) Burst loss probability versus load, and (b) Average end-to-end delay versus load, when the initiating nodes is source, center hop, and destination in the same network to provide differentiation through signaling.

References

1. C. Qiao *et al.*, "Optical Burst Switching - A New Paradigm for an Optical Internet," *JHSN*, *Vol. 8, No. 1*, pp. 69-84, 1999.
2. L. Xu *et al.*, "Techniques for Optical Packet Switching and Optical Burst Switching," *IEEE Communications*, *Vol. 39, No. 1*, January 2001, pp. 136-142.
3. M. Yoo *et al.*, "QoS Performance in IP over WDM Networks," *IEEE JSAC*, *Vol. 18, No. 10*, pp. 2062-2071, Oct. 2000.

TuJ3 **2:30 PM**

Limits of Effective Throughput of Optical Burst Switches Based on Semiconductor Optical Amplifiers

H. Buchta, E. Patzak, J. Saniter, *Heinrich-Hertz-Institut für Nachrichtentechnik Berlin GmbH, Berlin, Germany;* C. Gauger, *University of Stuttgart, Institute of Communication Networks and Computer Engineering, Stuttgart, Germany, Email: buchta@hhi.de.*

Both signal degradation and burst losses limit the effective throughput of optical burst switches. These limitations are analyzed for different burst switch architectures, bit rates, and numbers of fibers and wavelengths.

1. Introduction

Optical Burst Switching (OBS) is a promising candidate for a more dynamic optical layer to support the next generation Internet [1]. It can be considered a compromise between Optical Packet Switching (OPS) and Optical Circuit Switching. In an OBS network, edge nodes assemble several IP-packets with the same egress node and QoS class electronically into variable length optical bursts, which stay in the optical domain until they reach the egress edge node. Typical burst lengths are between a few µs and several 100 µs. Therefore switching times should be below 1 µs. Semiconductor optical amplifier (SOA) based switches with switching times in the ns range are well suited for this application. The maximum throughput of a node is limited by signal degradation caused by power loss, noise and crosstalk.

A key characteristic of OBS is the one-pass reservation scheme of network resources for each individual burst [1]. Bursts are sent without an acknowledgement of successful path setup and burst loss can occur in case of contention. The burst loss probability B can be reduced by using many wavelengths per fiber in combination with λ conversion and additionally by fiber delay lines (FDL) as buffers. For a given acceptable burst loss probability (B), node architecture and burst reservation scheme determine maximum utilization of WDM channels.

First we extend our scalability analysis of OBS nodes [2] by considering power loss, noise and crosstalk for nodes with FDLs and limited range λ converters. For nodes with 8 and 16 input/output fibers and for different line rates (2.5, 10, 40 Gbps) the maximum throughput is calculated, which is limited by the number of possible wavelength M. Second we determined the maximum utilization of WDM channels for the nodes for a burst loss probability of $B = 10^{-6}$. This gives us, the effective throughput, which is the product of maximum throughput and maximum utilization.

By considering both, physical constraints as well as results from performance evaluation of optical burst switches, a better understanding of the potential of future optical burst switching networks is achieved.

2. Node Architectures for Optical Burst Switching

To reduce cost and signal degradation we consider architectures with only one SOA in the signal path. For OPS, several similar one-stage architectures have been investigated [3, 4]. Fig. 1a shows a broadcast-and-select node architecture adapted for OBS which we call the tune-and-select (TAS) node [2]. This node has N input/output fibers and M wavelengths per fiber. It is strictly non-block-