On trains and wagons: switching variable length packets in a slotted OPS network

Chris Develder, Mario Pickavet, Piet Demeester
Dept. of Information Technology (INTEC), Ghent University – IMEC,
St-Pietersnieuwstr. 41, 9000 Gent, Belgium
email: {chris.develder, mario.pickavet, demeester}@intec.rug.ac.be

***Abstract*— Optical packet switching allows to fully and efficiently exploiting the capacities offered by (D)WDM. In this paper we investigate how slotted optical switches can deal with variable length packets, which is typical of IP.**

I. Introduction

To satisfy the ever lasting hunger for bandwidth, nowadays' communication networks resort to the deployment of (D)WDM networking. An initial step taken today is migration from still predominant point-to-point systems to real optical networking supporting circuit-switched optical paths [1]. Yet, despite their relative ease of design and operation, they suffer from the difficulty of dealing with highly variable traffic. Optical Packet Switching (OPS, [2]) is a longer term strategy exploiting fast optical switching techniques to offer better bandwidth granularity, efficiency and flexibility. The main difference with the Optical Burst Switching (OBS) concept [3], is that OPS operates in a slotted mode: packet arrivals at the inputs are aligned to slot boundaries and packets arriving in the same slot can be switched jointly.

Despite the essentially slotted concept, OPS switches can be used to deal with variable length packets by chopping them into chunks fitting within one slot. Thus, we obtain a train of slots constituting a single variable length packet. As indicated in Fig. 1, there are essentially two ways to treat these trains: either treat them as a whole and take decisions for the whole train at once, or rather treat each wagon (i.e. slot) independently.

A comparison of the train versus the wagon approach has been presented in [4] for shared optical busses using an access protocol for high-speed LANs/MANs. The authors discussed the overhead reduction attained by using a train-approach, and studied the delay vs. troughput behavior to conclude that for short train lengths the wagon approach proved to be more efficient.

In this paper we try to find out which approach is the best in terms of logical performance (ie. data loss due to contention, delay, service differentiation capabilities) for an optical packet switched WAN comprising slotted switches. In the next section II, we outline the switch architecture and the scheduling algorithm used. The subsequent Section III presents the simulation set-up taken to answer the train-or-wagons question. The results are summarized in Section IV, before concluding in Section V.

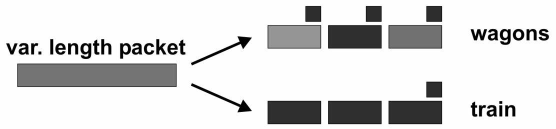

Fig. 1. Dealing with variable packets in a slotted switching concept: train versus wagons.

II. An Optical Packet Switch

The switch architecture we focus on was proposed within the European research project DAVID [6]. It consists of an all-optical switch matrix based on SOA technology. In- and output ports of this matrix are connected to F fibers (each carrying W wavelengths) providing connections to other switches. Wavelength convertors are provided to help solving contention. In addition, B ports are connected to a recirculating FDL buffer, which is fully shared among all I/O ports. The node structure is outlined in Fig. 2.

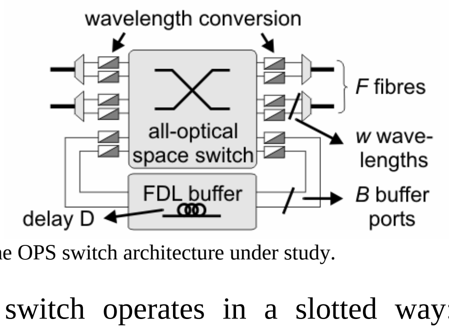

Fig. 2. The OPS switch architecture under study.

The switch operates in a slotted way: every timeslot, it inspects packets arriving at its input

This work has been supported by the European Commission through the IST-project DAVID (IST-1999-11387), and by the Flemish Government through the IWT GBOU-project "Optical Networking and Node Archi-tectures". C. Develder is supported as a Research Assistant of the Fund for Scientific Research – Flanders (F.W.O.–Vl.), Belgium.

ports, and subsequently decides what packets to forward or to drop. This decision is taken by following a fixed procedure, comprising two phases: (i) for each output fiber of the OPR, elect at most W packets to be forward directly, (ii) from the remaining packets, elect at most B to put in the buffer; any other packet will be lost. We do not consider deflection routing, since it is only effective at low network loads and can lead to out-of-order packet deliveries.

In this paper we investigate the behavior of such a single switch when the traffic it needs to forward consists of variable length packets. To handle these packets with the slotted switch, they need to be segmented into slots. One way to forward the resulting trains is to deal with each "wagon" individually. This implies that each slot needs to have an individual header, as indicated in Fig. 1.

The alternative, sometimes referred to as Slotted Variable Length Packets (SVLP) [7], is to treat the complete train as a whole and take a decision for the train upon arrival of its first wagon. Thus, a single full header (containing e.g. source, destination address, traffic class) suffices for the forwarding process. The fact that the following wagons belong to the same train can be indicated through e.g. a continuation bit field [4].

Note that we assume that the headers are transmitted on an orthogonal channel, e.g. through ASK/DPSK modulation [5]. Thus, the train length measured in slots will be the same for the train and wagon approach.

III. Simulation scenario

To compare the train versus wagon approaches, we focus on a single switch. To obtain the various performance parameters, we resorted to simulation. To guarantee trustworthy results, we used a high-quality random generator and assured 95% confidence on the results (for the sake of clarity, error margins are however not shown on the graphs). The parameters used are listed in Table I.

Table I. Simulation prameters.

Param.	Value	Meaning
F	6	nr. of I/O fibers
W	8	nr. of wavelengths per fiber
B	[0, 64]	nr. of recirculating buffer ports
D	2·L	FDL delay
L	[1.5, 20]	average train length (unit = slots)

A. *Traffic model*

The traffic model used for the simulations is clearly a packet train model, similar to the one proposed in [8]. For both the train lengths and gaps between successive trains we used a negative exponential distribution, while the inter-wagon gaps were always zero (cf. a train consists of wagons in successive slots). While this distribution may not be the most realistic one when trains are interpreted as being IP packets, the qualitative conclusions of our results are fairly independent of the train length distribution.

B. *Performance criteria*

The main performance criterion in an OPS environment is the loss rate: packets (trains) can get lost if both wavelength conversion and buffering fail to solve inevitable contention. Since a train is considered to be a single data unit, we assume it to be lost as soon as a single wagon is dropped at the switch. Important in the variable packet length concept is also the fairness of the scheduler: does it discriminate long trains against shorter ones?

Another performance criterion is delay. However, in the context of OPS WANs, it is only of secondary importance, since it will be limited compared to propagation delays and delays in access and metro parts of the network. Therefore (and because of space limitation), we will not discuss it in this paper.

As the OPS network will need to transport various traffic classes, it should be apt to support service differentiation. In [9], we investigated multiple service differentiation approaches for asynchronous variable length packets. In a slotted environment, the simplest approach is to use a simple priority mechanism: indicate the priority in the packet's header and give strict preference to higher priority packets when making the forwarding and buffering techniques. This was proven to be very effective in a fixed-length packet environment [10]. In this paper we investigate if it also is suitable for a train or wagon approach for variable length packets.

A last criterion is processing overhead. It is clear that since a wagon model requests every wagon to have its own header, the amount of forwarding decisions to be made at the switch will be a factor higher than in case of a train approach, roughly equal to the number of wagons per train.

IV. Trains or wagons

In this section we try to answer the question: should we adopt a train or a wagon approach? We first look at the loss rates for increasing loads in subsection A. The influence of the slot granularity (ie. ratio of slot length vs average train length) is investigated in part B. The last subsection C focuses on service differentiation capabilities of both train and wagon approaches.

A. *Influence of load*

Since the loss rate for a given load will clearly depend on the amount of buffer, we provide results for three sample buffer sizes: no buffer (B=0), four (B=4) and eight (B=8) recirculating buffer ports. The loss rates for this set-up are plotted in Fig. 3(a). Clearly, the loss increases with higher loads, and buffering aids in limiting the loss.

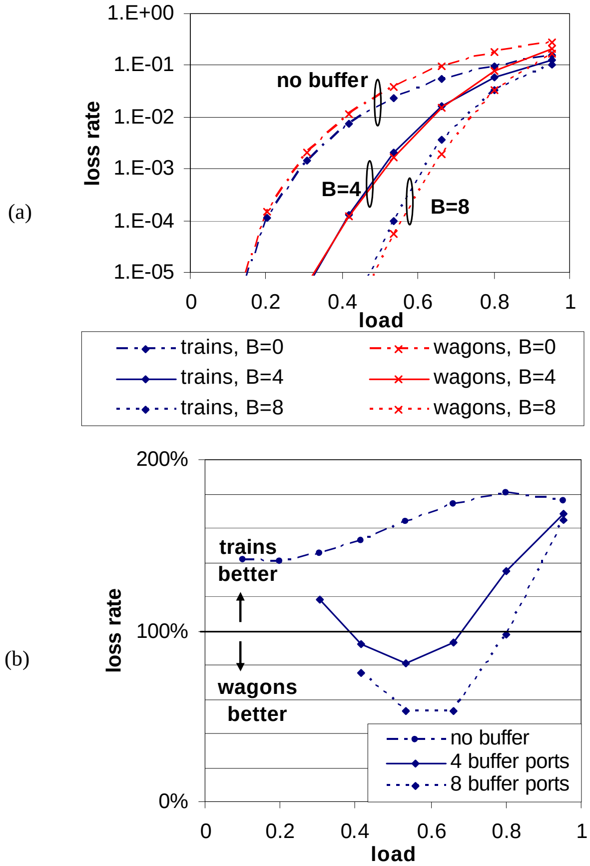

Fig. 3. Wagons vs trains: loss rates for increasing load, with B=0,4,8 buffer ports: (a) loss rate, (b) ratio of loss rate: wagons / trains.

Comparing the wagon versus train approach, we show the ratio of the loss rate attained by the wagon approach divided by that of the train approach in Fig. 3(b). When there is no buffer (B=0), we find that the wagon approach performs worse (ratio above 100%). However, when a buffer is present, the wagon approach achieves lower losses. Still, this is only the case for a particular load range: for very high or very low loads, the train approach performs better.

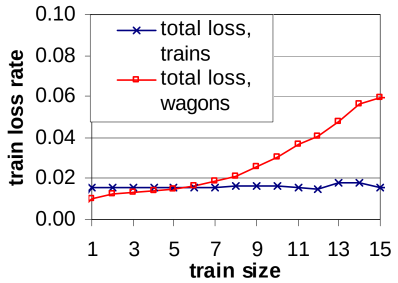

Fig. 4. Illustration of unfairness in loss rates for the wagon approach: loss rate per train length, for B=4 and a load of 0.62.

As indicated previously, loss rates may depend on the train length. In Fig. 4, we plot the loss rates per train size for B=4 buffer ports and a load of 0.62. Since the buffer size is chosen such that it can accommodate about 95% of the train lengths, the train scheduling approach is quite fair. For the wagon approach however, since each slot is treated independently and a train is lost as soon as a single wagon is dropped, the unfairness is quite severe (max vs min loss rate differ with about an order of magnitude).

B. *Influence of granularity*

The efficiency of handling variable length packets with a slotted switch will greatly depend on the slot resolution. For a given train size distribution, the choice of a given slot size will obviously determine the amount of wasted bandwidth because of padding. But even when this is ignored, the performance in terms of loss will also be influenced.

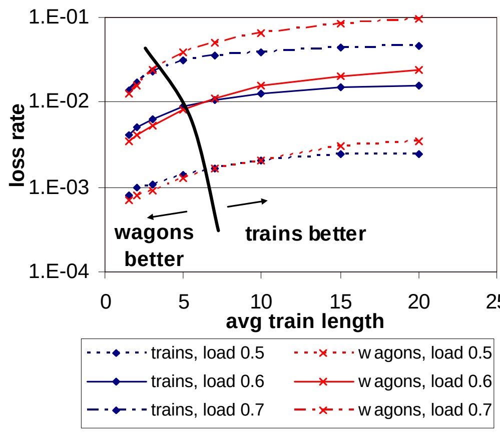

Fig. 5. Wagons vs trains: loss rates for increasing train lengths, with B=4 buffer ports, loads=0.5, 0.6, 0.7

In Fig. 5 we plot the loss rates for increasing average train length. As intuitively expected, the wagon approach only performs better for small train lengths. The crossover point moves slightly to larger train lengths when the buffer is increased.

C. Service differentiation

A simple priority mechanism, based on priority indicated in a packet's header showed to provide adequate class separation in a fixed packet length environment [10]. In this section we consider the same approach for trains an wagon approaches. As an example, Fig. 6 shows the loss rates when using two priority classes, with 40% of the traffic having the high priority. Since in the train approach, high priority packets cannot preempt lower priority trains that arrived a few slots earlier, the differentiation achieved is far less pronounced compared to the wagon approach. The low priority loss rates dominate the overall loss rate, which evolves as in the priority-less case as plotted earlier in Fig. 3(a).

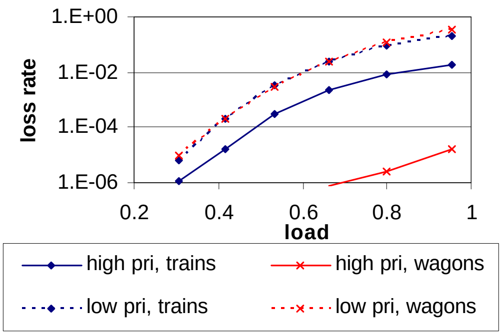

Fig. 6. Service differentiation of wagons vs trains: loss rates for increasing loads, with B=4 buffer ports.

V. Conclusions

To our knowledge, this is the first paper to quantitatively compare the train vs wagons approach for optical packet switches in a WAN context, ie. a mesh of optical packet switches rather than a MAN/LAN environment with an access protocol. In a slotted OPS switch variable length packets, which are splitted into slots, can be treated either as a whole (trains) or on a slot-by-slot basis (wagons).

The wagon approach can help to reach lower overall (train) loss rates when there is a buffer, and trains are relatively short (ie. a few slots). The more buffer, and the shorter the trains, the greater the potential advantage is. However, this only holds for a limited range of loads: when the load is either low or rather high (order 0.8 and above), the train approach is to be preferred.

From a service differentiation point of view, the wagon approach is able to reach more pronounced service differentiation when a simple priority-based approach is adopted.

The potential advantages of a wagon approach are paid for by an increased control overhead and load on the scheduler (factor of order of average train length measured in slots) and unfairness, in the sense that it more severely discriminates longer trains.

References

[1] B. Mukherjee, “WDM optical communication networks: Progress and challenges”, JSAC, vol. 18, no. 10, Oct. 2000, pp. 1810–1824.

[2] D.K. Hunter, I. Andonovic, “Approaches to optical Internet packet switching”, IEEE Comm. Mag., vol. 38, no. 9, Sep. 2000, pp. 116–122.

[3] C. Qiao, M. Yoo, "Optical Burst Switching - A new paradigm for an optical Internet", J. High Speed Networks, vol. 8, no. 1, Jan. 1999, pp. 69–84.

[4] B. Mukherjee, A.E. Kamal, “Scheduling variable-length messages on slotted, high-speed fiber optic LANs/MANs using the continuation-bit approach”, Proc. INFOCOM 1991, vol. 2, pp. 678–687.

[5] N. Chi, et al., “Optical Label Swapping and Packet Transmission Based on ASK/DPSK Orthogonal Modulation Format in IP-over-WDM Networks”, Proc. OFC2003, paper FS2, vol. 2, pp. 792–793.

[6] L. Dittmann (ed.), et al., “The European IST Project DAVID: a Viable Approach towards Optical Packet Switching”, to appear in JSAC.

[7] F. Callegati, et al., “Assessment of packet loss for an optical feedback buffer node using slotted variable-length packets and heavy-tailed traffic”, Proc. ICTON 2002, vol. 1, pp. 51-56.

[8] R. Jain, S. Routhier, “Packet Trains — Measurements and a New Model for Computer Network Traffic”, JSAC, vol.4, no. 6, sep. 1986, pp. 986–995.

[9] C. Develder, J. Cheyns, M. Pickavet, and P. Demeester, “Service differentiation for OBS in a switch with recirculating FDL buffer”, in press.

[10] C. Develder, M. Pickavet, and P. Demeester, “Strategies for an FDL based feed-back buffer for an optical packet switch with QoS differentiation”, Proc. COIN 2002, paper COIN.TuD1, pp. 114–116.