Performance of the Just-Enough-Time (JET) scheme for Optical Burst Switching

A. Myers, P. Bayvel
Optical Networks Group, Dept. of Electronic & Electrical Engineering,
UCL (University College London)

Abstract: In this paper the Just-Enough-Time (JET) optical burst switching mechanism is investigated. The role of two important parameters, the per-hop-offset and burst-length/processing-time ratio are examined. Results are presented on three network topologies to demonstrate that faster processing at each node leads to higher efficiency, and that a significant decrease in data loss can be achieved by increasing launch offset-times. However, even for modest traffic levels data loss rates are high. Some numerical approximations to the loss rate that can be used for network dimensioning are described.

1. Introduction

Optical burst switching (OBS) has been proposed as a method of operating a fast packet data network avoiding optical-electrical-optical conversion or (currently unavailable) optical buffering at intermediate nodes within the network. Two mechanisms for OBS have been suggested. Schemes similar to that in [1] operate on circuit-switching principles, with circuits set up for the short duration to transmit an amount of pre-buffered data, In contrast, **t**he JET scheme [2] operates without guaranteed packet delivery but has simpler control mechanisms. An alternative to OBS entirely is the Static Wavelength Routed Optical Network (S-WRON) with permanent wavelength paths established between all pairs of nodes, however this approach lacks flexibility in the face of rapidly changing traffic patterns. Previous studies of JET have examined only the performance of a single node [3 or only regular networks [2]. This work investigates the performance of JET operating on networks dimensioned so that an S-WRON is an alternative method of data delivery. An important contribution of this work is to suggest where JET might or might not be a suitable alternative to the S-WRON.

2. 2. Description of JET

In JET, network nodes play two roles. In the first role, incoming data from end-stations is buffered according to its destination. After some time the data is ready for dispatch as an optical burst. A data signal (the burst header) is then sent to the next downstream node and some time later - $t_{offset(launch)}$ - the burst is transmitted on the wavelength specified in the header. t_{offset} is the time delay between a header and its repspective data, and is sufficient for the downstream node to fulfill the second role: The arrival of a burst header on the control channel of a link signals a node to attempt to reserve a wavelength/time-slot for the soon-to-arrive data to be switched to an output link closer to the destination. Full wavelength translation capability at each link is needed so that any burst can be routed to any free wavelength on the output link; therefore the wavelength of a burst has local significance only. The downstream node then sends a new header to the next downstream node. At each hop t_{offset} is reduced (Fig.1) by the processing time (per-hop-offset or τ_{pho}) at each node; therefore for a burst to travel n hops, $t_{offset(launch)} \geq n\tau_{pho}$. The advance notice provided by the header suffices that when the data-burst arrives at an intermediate node, that node is already set to route the signal from input to output channel to output channel.

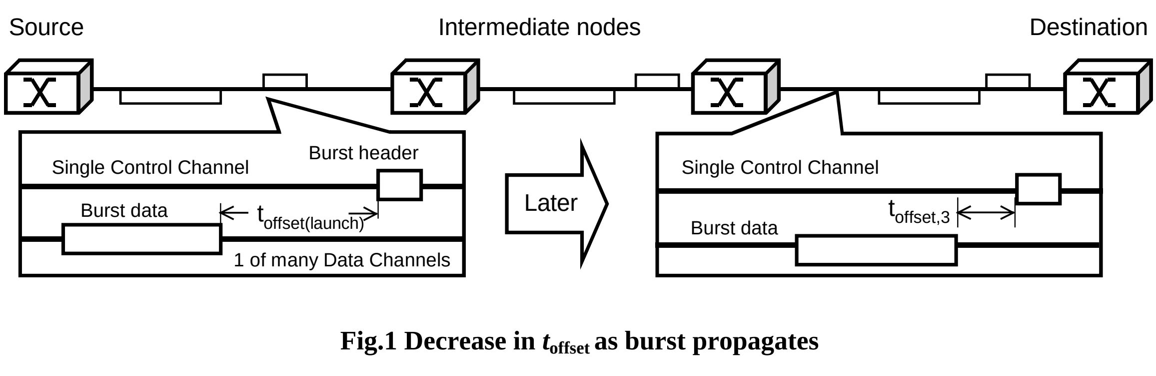
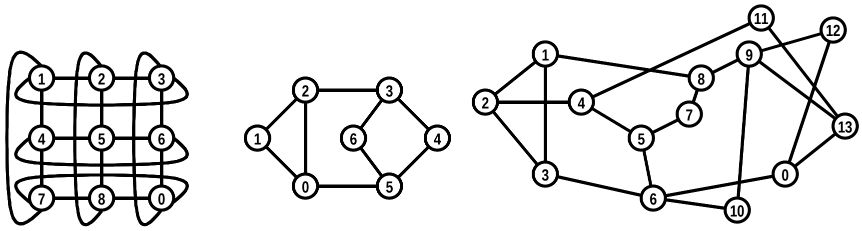


Fig.1 Decrease in t_{offset} as burst propagates

3. Simulation Details

In this work, experimental results are presented that were obtained from a C++ simulation which **tracks** the path and timing of individual burst headers as they progress through the network. For every pair {source,destination} of nodes a sequence of data bursts (a *flow*) was generated as a Poisson process with a given mean rate ν. This distribution was chosen for simplicity and to allow a more analytical approach to be used in regions where simulation is impossible. A routing table specifying the next exit-link for each burst based on {source,destination,current-link} - calculated off line as described in the next section - was used to generate a sequence of reservation requests for each burst. By keeping track of the usage of each channel on each link at every time-value it is possible to determine which requests could be honoured, and which bursts would be lost as unrouteable. The proportion of lost bursts of a flow is equated with the burst-loss probability for that flow.

The network topologies illustrated in this paper are shown in fig.2. The aim of using three distinct topologies is to show that certain parameters have a significant impact on performance irrespective of the details of the topology. Each link shown is bi-directional and all have a length (propagation time) = 100 arbitrary units.



(a) TOR3 3 channels | **(b) ARMNet** 6 channels | **(c) NSFNet** 13 channels

Fig.2 - Experimental Toplogies

The number of wavelengths in the network (each wavelength being an individual output channel on every link) is chosen to be the same as is needed on the most congested link if an S-WRON were to be created over the same topology. This is only three channels for the small, well-connected TOR3 graph but 13 on the sparser NSFNet topology.

In [4] we have shown that the loss probability for variable length bursts has a strong positive correlation with the burst length – long bursts are hard to accommodate. In order to draw a correspondence between burst loss probability and data loss rate bursts of fixed lengths are assumed (T_{burst} = 100 arbitrary time units).

3.1. Flow intensity

The traffic pattern is assumed to be an equal intensity flow of bursts between every pair of nodes, with mean arrival rate $\nu = uB/T_{burst}$ where B is the bandwidth of a single wavelength channel (nominally 40Gbit/**s**) and the utilization u is the fraction of a single wavelength bandwidth occupied by the data between every pair of nodes; deliberately described as a unit-less quantity varying between 0 and 1.1 to remove direct dependence of the results on B. For each given topology an S-WRON could carry without loss traffic up to u=1.0 between each s-d pair with the number of wavelengths specified.

3.2. Routing table derivation

For each {*source node, destination node, current node*} tuple for potential arriving burst headers, a list of first and second choice next-hops is given in advance to the simulation. At the time of a header's arrival at a node, the model attempts to reserve a channel on the links to one of each of the first-choice nodes, selected in random order. If no such link has a channel unreserved for the duration of the burst-data then a set of second-choice links is searched, also in a random order. For each potential next-hop a feasibility condition is imposed, that there must be at least one path from this next-hop to the destination for the burst to arrive with $t_{offset} \geq 0$. If no feasible first or second choice link is available the burst is dropped. We will consider two methods of routing table allocation:

In the first, *multiple-shortest-path* (MSP), the first-choice next-hops are the nodes on a shortest-path route to the destination. All other nodes are considered as potential second-choice next-hops, subject to the feasibility condition. The traffic between source and destination spreads across multiple links and can take a variety of paths, leading to a variation in latency, and the possibility of out-of-order burst arrivals.

In the second, *designated routing* (DR), only a single first-choice of next-hop is permitted, and no second-choices. All bursts in a flow are forced to follow a fixed path and are guaranteed to be delivered in order if at all. The set of paths is chosen to minimize the congestion at each link according to an algorithm not described here.

3.3. Per-Hop Offset time (τ_{pho})

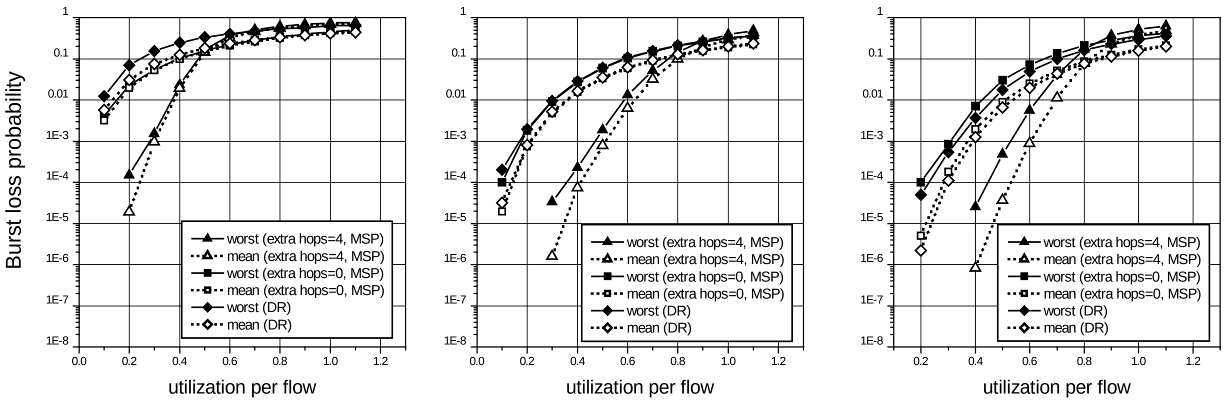
A significant physical parameter of the network is the amount by which t_{offset} decreases for a burst as it traverses consecutive nodes – τ_{pho}. This is the processing time a node requires to do tasks such as header decoding, routing look-up, output channel scheduling and transmission of a new header to the next node. We will consider how the loss-rate of the network varies as τ_{pho} is varied from 0.01 T_{burst} to $3T_{burst}$.

3.4. $t_{\text{offset (launch)}}$

The offset time that the initial node sets between the header and data of a burst determines how many hops through the network the burst can make. In order to allow the burst to follow a path longer than the shortest between the source and destination we increase $t_{\text{offset(launch)}}$. If the shortest path between nodes {i,j} is n_{ij} hops then results will be shown for $t_{\text{offset (launch)}} = \tau_{\text{pho}}(n_{ij}+ 4)$ allowing for 4 extra hops. In [5,6] similar schemes under the headings of "deflection routing" or "hot-potato routing" are considered, where bursts divert from their optimal route depending on conditions – however using complex synchronisation schemes, buffers and delay-lines.

4. Simulation Results and Discussion

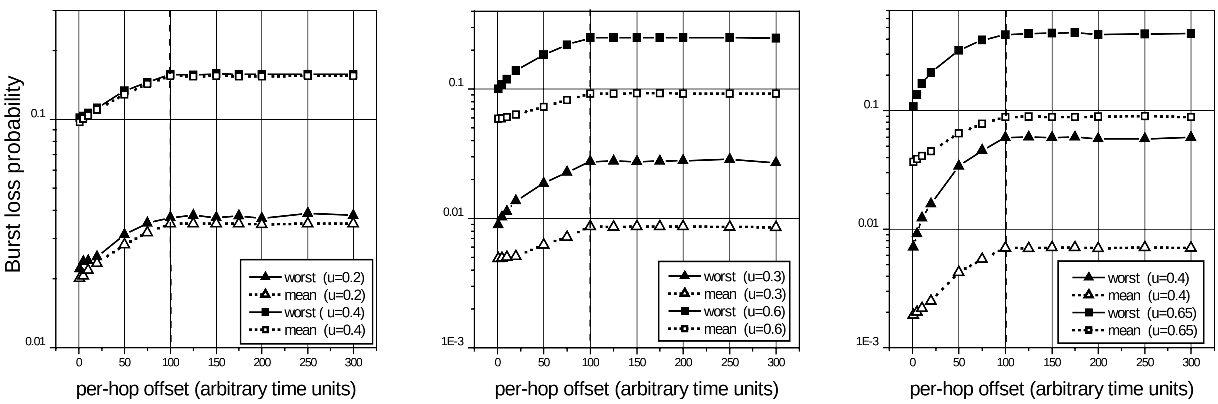
Fig.3 Shows the mean burst loss probability averaged over all flows in each topology, alongside the result for the lossiest flow at each utilization value.



(a) TOR3 (b) ARMNet (c) NSFNet

Fig.3: burst loss probability vs. utilization
τ_{pho} is 1 time unit. Simulation duration is 10^7 time units.
Routing is DR, MSP, and MSP with t_{offset} increased to permit 4 extra hops for every burst.

The results demonstrate a clear advantage to increasing the number of hops a burst may traverse by reducing the probability of its loss, although inevitably at the expense of an increase in latency. This concurs with the results of [6]. The results show choice of route determination is less significant. In the ARMNet graph there are, for many {source,destination} pairs, more than one path of equal shortest length; MSP allows traffic to spread over these paths according to channel availability at the time of scheduling, while DR allows only a single next-hop for each flow and therefore less potential output channels from each node. However, in NSFNet DR performs better because it is able to more evenly distribute traffic to avoid "hot-spots" of congestion.



(a) TOR3 (b) ARMNet (c) NSFNet

Fig. 4: burst loss probability for selected utilizations, with varying per-hop offsets
.Simulation duration is 10^7 time units. Routing is MSP. Burst length is 100 time units. No extra hops allowed

Fig.4 shows the effect of increased τ_{pho}. The network becomes less efficient as the range of offsets with which bursts arrive at a node increase. A large per-hop-offset means that a wide range of offsets exist in the system at any time; this leads to a decrease in the efficiency of the use of each channel and hence an increased loss rate. In [2] the authors suggest that a useful element of QoS provision could be included in a JET system by providing higher priority bursts with greatly increased $t_{offset(launch)}$. We have shown that this will lead to a decreased overall efficiency of the network, since this method of prioritisation is not work-conserving. Note that loss probability ceases to rise once the per-hop offset increases to equal the burst length (100 time units) beyond which point bursts are completely separated into strict priority classes at each node according to their value of t_{offset}/τ_{pho}.

5. Numerical model for loss rate, with designated routing

In the case of DR routing with low τ_{pho}/T_{burst} with k channels per link a good heuristic approximation for the loss rate of a flow q entering node p from link i with exit-link j is given by $p_{loss} = \frac{U^k/k!}{\sum_0^k U^n/n!} - \frac{R^k/k!}{\sum_0^k R^n/n!}$ where U is the sum utilization of flows with next-link i, and R is the sum utilization of flows with entry-link i *and* next-link j. The first term in the numerator is the Erlang-B formula for the loss rate at an M/D/k/k queue where arrivals are uncorrelated. The second term in the numerator is a correction for the correlation between arrivals on the same link. Since the output of one link becomes the input to the next, This formula is used iteratively throughout the network topology to obtain a converged solution for an estimate of the loss rate for each flow. Results are shown compared to the per-burst simulation in fig. 5(a), showing excellent agreement for all values of utilization.

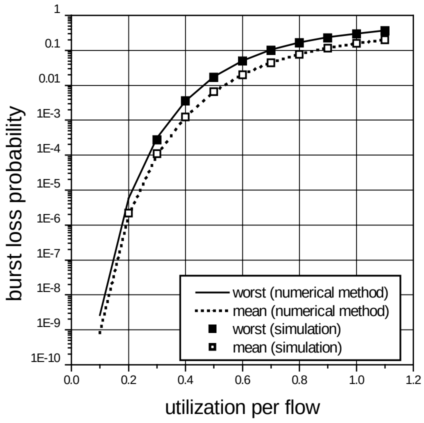


Fig.5(a) NSFNet: loss rate vs. utilization 13 channels

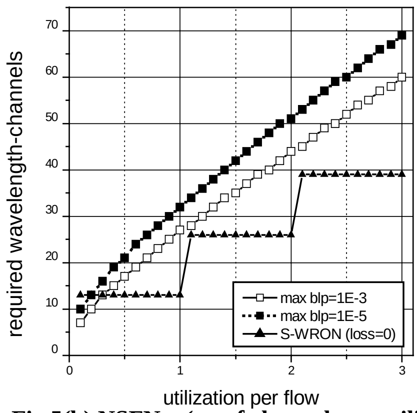


Fig.5(b) NSFNet (no of channels vs. utilization) JET compared to Static Wavelength Routed Optical Network

This approach is used to determine the number of channels required for all flows to have losses lower than any desired bound (chose for illustration as 10^{-3} and 10^{-5}) by repeating the calculation for increasing numbers of channels until the bound is met. Fig. 5(b) shows how many wavelengths are required for DR-based JET on NSFNet for these two values of loss probability as the traffic load increases. For comparison we show the number of wavelengths required on the most congested link for an S-WRON that carries the same level of traffic – an integer multiple of 13. The only situation where JET is better is for very low loads ($u < 0.3$) where an S-WRON requires at least a certain minimum number of channels to avoid wavelength contention.

6. Conclusion

In this work we have shown that τ_{pho}/T_{burst} must be minimized to improve the efficiency of the JET network and that increasing $t_{offset(launch)}/\tau_{pho}$ to allow more hops for each burst makes a significant improvement. Considering that JET requires the complexity of full wavelength translation at every node, and under strict requirements for ordered burst delivery, JET has clear advantages over S-WRON only at loads < 0.3 between each source-destination pair of nodes**.** For higher loads, a network architecture that avoids random burst collisions must be employed for better efficiency.

The financial support of the EPSRC and of Nortel Networks is gratefully acknowledged. M. Düser, A. Zapata and I. de Miguel are thanked for their assistance during this work.

[1] M. Düser, E. Kozlovski, R. I. Killey and P. Bayvel, "*Design trade-offs in optical burst switched networks with dynamic wavelength allocation*",Proc. ECOC 2000, vol 2 pp23-24
[2] C.M. Qiao , M.S. Yoo , "Optical burst switching (OBS) - a new paradigm for an optical Internet," *J. of High Speed Networks* 8 (1): 69-84 1999
[3] C. Gauger, K.Dolzer, J. Spath, S. Bodamer, "*Evaluation of Reservation Mechanisms for Optical Burst Switching*". International Journal of Electronics and Communications (AE) 55, (1) 2001
[4] A. Myers, "*JET performance: interim report*" – internal ONG report.
[5] A. Bononi, G. Castanon, O. Tonguz, "*Analysis of Hot-Potato Optical Networks with Wavelength Conversion,*" J. Lightwave Technology, 17 (4) 525-533 1999
[6] G. Castanon, L. Tancevski, L. Tamil, "*Optical packet switching with multiple path routing,*" Computer Networks 32 (2000) 653-662