

A Comparative Study of OBS and OFS

Chunsheng Xin and Chunming Qiao

Department of Computer Science and Engineering, SUNY at Buffalo

TEL: 716-645-3180x140, FAX: 716-645-3464, Email: xin@cse.buffalo.edu, qiao@computer.org}

Abstract: This paper evaluates the performance of two promising approaches for IP over WDM networks, namely optical burst switching (OBS) and optical flow switching (OFS), based on both synthetic and real traffic.

I INTRODUCTION

OBS [1,2] and OFS [3] are two new optical switching schemes for WDM networks. In this paper, we consider a core WDM mesh network consisting of entirely OBS nodes (similar to packet switches) or entirely OFS nodes (similar to cross-connects used for circuit switching), each of which is attached with several edge (ingress or egress) routers via proper interface units. We assume that IP-based routing and MPLS-like signaling protocols are used to control the WDM network (whether it is burst-switched or circuit-switched).

A. Optical Burst Switching

In OBS, each ingress router (more accurately, the interface unit between it and the WDM core, here after, we will omit the phrase "interface unit") assembles incoming data (IP) packets going to the same egress router into a burst according to some burst assembly schemes. For each burst, a control packet is first sent out on a control wavelength to the egress router and the burst will follow (on a separate data wavelength) after an offset time, (this offset time can be made to be no less than the total processing delay to be encountered by the control packet in order to eliminate the need for FDLs and at the same time much less than the round-trip propagation delay between the ingress and egress routers). The control packet, which goes through O-E-O conversion at every intermediate node, attempts to reserve a data wavelength (for just enough time to accommodate the following burst) on the outgoing link, and, if the reservation succeeds, configure the optical switching fabric to switch the following burst. If the reservation fails (because there is no wavelengths available at the time of the burst arrival on the outgoing link), the burst will be dropped (and may be retransmitted by a higher layer protocol such as TCP). If a burst arrives at the egress router, it will be dis-assembled into data packets. In this study, we consider the burst assembly scheme whereby packets (going to the same egress router) that arrived during a fixed period, called the burst assembly time (BAT), are assembled into a single burst. Packets arriving after the next assembly cycle begins (i.e., when a timer initially set to BAT units expires) will be assembled into a different burst (even though they may go to the same egress router as the packets contained in the previous burst). Clearly a smaller BAT value will likely result in shorter bursts, and thus higher amortized overhead introduced by control packets (and guard bands). On the other hand, a larger BAT value will likely result in a longer end-to-end (E2E) delay of all but one packets in each burst. Note that this scheme is simple to implement and results in a bounded (i.e. maximum) assembly delay. However, one may use other burst assembly schemes to guarantee a minimum burst length to reduce the amortized overhead at the expense of having an unbounded assembly delay or try to capture an upper-level protocol data unit (PDU), such as a TCP segment, into a burst.

B. Optical Flow Switching

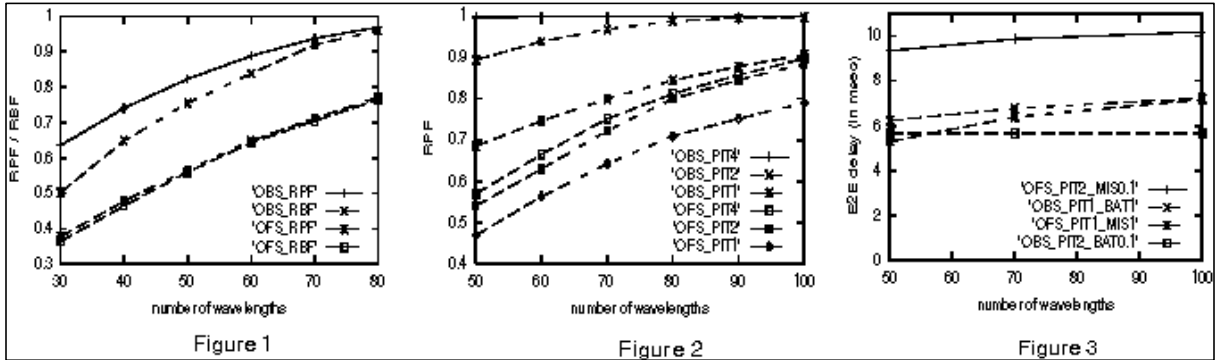
A flow may be defined as a sequence of consecutive packets from the same ingress router to the same egress router. One approach to OFS is for an ingress router to detect the beginning (and end) of a flow and request for the set-up (and release, respectively) of a lightpath (or wavelength path) for this flow. A lightpath occupies a λ on each link along the path from the ingress to the egress router. Other approaches to OFS (including one that exercises a network-wide control over the set-up and release of all lightpaths) are also possible but not considered in this paper. For the purpose of this comparative study of OBS and OFS, we assume that, when a packet arrives, the ingress router simply checks if there is an existing flow to which the packet belongs. If yes, this packet is either sent immediately over the established lightpath for that flow or be buffered if such a lightpath has been requested and is being set up. If not, the packet is considered to be the first packet of a new

flow and is buffered. A lightpath set-up request is then sent to establish a lightpath to the egress router in a hop-by-hop fashion for this packet (and following packets belonging to the same flow). Buffered packets of a flow will be discarded when a negative acknowledgment (NAK) arrives at the ingress router informing that a lightpath for the flow cannot be established due to lack of available λ , or sent one-by-one after the lightpath is established (i.e. after a positive ACK arrives). Note that an ingress router considers that a flow ends if there is no packet going to the same egress router as this flow arriving within a period called maximum inter-packet separation (MIS) since the last packet of this flow arrived. As soon as a flow ends and the last packet of the flow is sent, a lightpath release request is sent to tear down the lightpath for that flow (again in a hop-by-hop fashion). One of the reasons we consider the above-mentioned OFS approach in this study is that it is the most similar to the OBS approach described earlier (i.e., with the BAT-based scheme). In fact, similar to BAT, MIS will also affect the performance of OFS as follows. A smaller MIS results in shorter flows which in turn result in more frequent lightpath set-ups and releases, and thus higher amortized overhead (in terms of packet delay, bandwidth waste during the set-up/release period, and processing load). On the other hand, a larger MIS results in longer gaps (idle time) between packets in a flow and thus a lower bandwidth utilization. Note that one may also detect the beginning and end of a flow based on some threshold values specifying the minimum number of packets (or bits) passing through an ingress node during a fixed period of time, or with the help of some upper-level protocols.

II PERFORMANCE COMPARISON

We have simulated the performance of OBS and OFS in a 10-node network with up to 100 data wavelengths (λ) per link and wavelength (λ) conversion at every node. The average number of hops between two nodes in the network (which is similar to the NSFnet and vBNS backbone) is only about 1.5 and the per hop propagation delay is assumed to be around 4 msec. It is assumed that one control λ is used for sending control packets in OBS or the set-up/release/ACK/NAK packets in OFS. We have used the network traces (packet streams) collected by NLANR [4] from the vBNS backbone. Since the traces are based on OC-3 link speed, we have scaled up the packet arrival and transmission rate to match the maximum link speed of 1 Tbps (which is achievable with 100 wavelengths each at OC-192). To evaluate the effect of varying load and other parameters which cannot be easily controlled with the traces, we have also generated the synthetic traffic that mimics real traffic in terms of packet length distribution (e.g., 20% are 0 or 1 byte packets, 23% are 1500-byte packets, etc.). In addition, the synthetic traffic has non-uniform distribution of ingress and egress routers (i.e. some nodes send/receive more packets than others). It is also asymmetric (i.e. a node may send more packets than it receives). The performance metrics we have used include the (averaged) received packet/bit fraction (RPF/RBF) (which is equal to 1 minus the dropped packet/bit percentage) and E2E delay (which includes the pre-transmission delay due to burst assembly in OBS or buffering of the packets in OFS at an ingress router). The following results have been obtained with the default values of BAT and MIS set to about 16 μ sec and 1.6 μ sec respectively. These default values yielded good performance for OBS and OFS, respectively (in fact, better performance than that obtained using some other values of BAT and MIS) for the traffic simulated.

Figure 1 compares the RPF/RBF of OBS and OFS under the vBNS traces as a function of the number of λ 's. We can see that OBS, which allow for statistical sharing of the bandwidth of a wavelength among bursts (belonging to possibly different flows according to the definition used for OFS), achieves a 50% higher RPF and 30% higher RBF than OFS. The reason for having a lower RBF than RPF in OBS is that the burst loss is not totally random. A longer burst (having more bits) is more likely to get dropped when the probability of burst collision is high (due to a small number of λ 's) and such a burst loss has more adverse effect on RBF than RPF. On the other hand, there is no significant difference between RPF and RBF in OFS because packets (long or short) belonging to a flow is a single switching unit, and hence the probability of a packet being dropped does not depend on its length. Figure 2 compares the RPF of OBS and OFS under the synthetic traffic as a function of traffic load which is represented by the (average) packet interarrival time (PIT), as well as the number of λ 's. Here, "PIT1" denotes the minimum PIT (i.e., the maximum network load



used in our simulation), and "PIT2" is twice of PIT1 (i.e., 50% of the maximum load simulated), and so on. We can see that, under either a heavy or a light network load, OBS achieves a substantially higher RPF than OFS (especially under light load with a small number of λ 's). However, when the number of λ 's increases, the difference between the performance of OBS and the performance of OFS reduces, implying that OFS can benefit from the increase in the number of λ 's more than OBS (to put a negative spin on the matter, one can say that, when the number of λ 's decreases, the adverse effect on the performance is more severe in OFS than in OBS, especially under light traffic load). This is because the basic channel granularity is one single λ which is provisioned at the time scale of a flow in OFS, whereas in OBS, the bandwidth of a single λ can be provisioned as a much smaller time scale (i.e. that of a burst), which results in a higher degree of statistical sharing of the bandwidth. Figure 3 compares the average E2E delay of OBS and OFS. Both "MIS1" and "BAT1" in the Figure denote the default values, and "MIS0.1" and "BAT0.1" are one tenth of MIS1 and BAT1 respectively. We can see that, with a small MIS (MIS0.1), OFS has a large E2E delay. This confirms our intuition that using a small MIS results in many short duration flows, each of which encounters a large lightpath set-up delay (which is no less than the round trip propagation delay). OFS can achieve a small E2E delay if most flows have a long duration so the set-up delay is amortized among many packets (but with a large MIS, it risks a low bandwidth utilization). On the other hand, the effect of BAT on the E2E delay in OBS is not significant. This is because, generally speaking, the BAT is several order of magnitudes smaller than even the one-way propagation delay, which dominates the E2E delay in OBS. The E2E delay comparison under the vBNS traces has a similar result.

III CONCLUSION

We have used simulations to compare the performance of two new optical switching schemes, OBS and OFS. Our results have shown that OBS achieves a significantly better throughput than OFS under the assumptions about the network and traffic that have been made. We believe that this is a result of more efficient just-enough-time bandwidth reservation scheme used in OBS. We have also found that the E2E delay in OFS depends very much on the MIS value used to detect the beginning/end of a flow. With a sufficient large MIS, the E2E delay in OFS can be lower than that in OBS. However, the throughput of OFS will suffer unless there is a sufficient number of λ 's (e.g. so much so that every flow can have a dedicated λ).

IV REFERENCE

1. M. Yoo and C. Qiao, "Just-enough-time (JET): a high speed protocol for bursty traffic in optical networks", IEEE/LEOS Technology for a Global Information Infrastructure, Aug 1997
2. C. Qiao and M. Yoo. "Optical Burst Switching (OBS) - A New Paradigm for an Optical Internet", a special issue of Journal of High Speed Networks (JHSN) on WDM Networks, Vol. 8, No. 1, pp. 69-84, 1999
3. Eytan Modiano. "Optical Flow Switching in the Next Generation Internet", Presented at the GBN 2000
4. National Laboratory Applied Network Research, <http://www.nlanr.net>