Architecture of a Hierarchical Time-Sliced Optical Burst Switching System

Vijay Sivaraman and Arun Vishwanath
School of Electrical Engineering and Telecommunications
University of New South Wales
Sydney, NSW 2052, Australia
Emails: {vijay@, arunv@ee.}unsw.edu.au

***Abstract*— To overcome the need for large buffers to store contending bursts in optical burst switched (OBS) networks, a recent variant called time-sliced OBS (TSOBS) suggested that bursts be sliced and spread across multiple frames of fixed-length time-slots. This paper generalises TSOBS to allow a hierarchy of frames. Termed hierarchical TSOBS (HiTSOBS), this scheme supports several granularities of rates, and permits multiple traffic classes with different loss-delay requirements to efficiently share the network. We present an architecture for HiTSOBS, and offer it as a viable option for the realisation of flexible and cost-effective OBS networks.**

I. INTRODUCTION

Optical Burst Switching (OBS) [1] is a hybrid of circuit and packet switching: aggregates of packets, called bursts, are switched atomically within the network, while a control packet is sent ahead of the burst to set up a short-lived end-to-end circuit for the burst. OBS thus combines the scalability of optics for fast data plane switching with the flexibility of electronics for switching decision control. An unfortunate consequence of this architecture is that the optical buffering required for contention resolution grows in proportion to the burst size. The control-plane advantage of large bursts is thus tempered by the larger buffers required in the data-plane.

A variant called Time Sliced OBS (TSOBS) was proposed in [2] to overcome this problem. Time is divided into frames that contain a given number of fixed-length slots. TSOBS slices a burst, and transports successive slices in the same slot location of successive frames. This preserves the control-plane scalability of OBS (since only one switching decision is required to switch all slices belonging to a burst), while drastically reducing the optical buffering required at switching nodes (since a contending burst need only be buffered a slice at a time, independent of burst size). Several architectures for implementing the optical time-slot interchanger (OTSI), a key component of the TSOBS system, are proposed and analysed in [2].

While TSOBS successfully addresses the scalability of optical burst switching systems, it is excessively rigid in its frame structure. The frame size (i.e. number of slots per frame) is a key parameter that has to be universally pre-configured at all switches. A small frame size increases contention probability since overlapping bursts are more likely to pick the same slot number, while large frame sizes induce larger end-to-end delays due to each flow having access to a reduced fraction of the link capacity (one slot per frame) leading to significant queueing delay at the ingress edge node. This loss-delay trade-off, determined by frame size, is uniform across all traffic flows, and cannot be dynamically adjusted, making TSOBS too rigid for practical use.

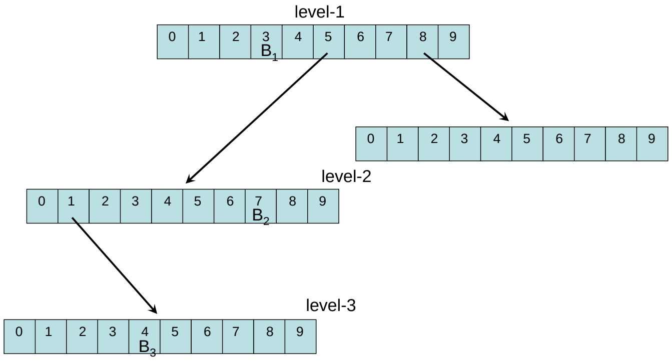

Fig. 1. HiTSOBS frame hierarchy

We overcome these limitations of TSOBS by generalising the frame structure to a flexible hierarchy. Our idea draws inspiration from the hierarchical round-robin (HRR) packet scheduler proposed in [3], and is termed hierarchical TSOBS (HiTSOBS). As we will elaborate in the following section, HiTSOBS allows multiple frame sizes to concurrently co-exist, with slots lower in the hierarchy progressively offering lower rate service. This allows delay-sensitive traffic classes to operate at higher levels of the hierarchy while concurrently supporting loss-sensitive traffic at the lower levels. Along with the ability to support differentiated services to different traffic classes, HiTSOBS dynamically adapts the frame hierarchy as the traffic mix changes, thus obviating network-wide pre-configuration.

II. ARCHITECTURE

A. Frame Hierarchy

Assume that time-slots are numbered consecutively, starting at 0. We select radix r which defines the number of slots in each frame in the HiTSOBS hierarchy. The top-level (level-1) frame therefore repeats every r slots. A burst transmitted at this level would occupy time-slots k, $k+r$, $k+2r$, …, $k+(B-1)r$ where k is the time-slot at which burst transmission starts and B is the size of the burst in slot units. For example, for radix $r = 10$, the burst B_1 of size 22 slots, shown in Fig. 1 to occupy the 3-rd slot in the level-1 frame, may be transmitted over time-slots 8043, 8053, 8063, …, 8253. Note that a given flow of bursts transmitted at level-1 has access to $1/r$ of the link capacity.

A slot in the level-1 frame may expand into an entire level-2 frame. For example, the 5-th slot in the level-1 frame in Fig. 1

expands into a level-2 frame. Successive slots in this level-2 frame are served in each successive turn of the 5-th slot of the level-1 frame. The burst B_2, shown to occupy the 7-th slot in this level-2 frame, may therefore be transmitted in time-slots 8175, 8275, 8375, and so on. Note that a burst transmitted at level-2 therefore has access to $1/r^2$ of the link bandwidth. Consequently, we can expect flows transmitting their bursts at level-2 of the hierarchy to have larger queueing delay at the edge compared to flows at level-1. However, the larger spacing between burst slices leaves more room for contention resolution using small optical buffers, making the losses for level-2 flows lower than that for level-1 flows.

The reader can extend the above structure to more levels; in general a slot in a level-i frame transports the burst at $1/r^i$ of the link capacity. It is also easy to map a time-slot number to its position in the frame hierarchy: the k-th digit of the time-slot number read backwards denotes its position in the level-k frame, and the process terminates when a leaf node is encountered. Returning to our example with radix $r = 10$ illustrated in Fig. 1, if we are asked to determine the contents of time-slot 8415, we would traverse the 5-th slot in the level-1 frame, the 1-st slot in the level-2 frame, leading to the 4-th slot in the level-3 frame, which is a leaf showing that a slice of burst B_3 is carried in that slot. Such an operation will be required for the control pane operation described next.

B. Control Plane Operation

The HiTSOBS ingress edge node accumulates data into bursts, and classifies them into an appropriate QoS class. For illustration purposes, say there are two classes: real-time traffic that needs low latency and is not very sensitive to loss, and TCP traffic that is not very sensitive to latency but requires low loss. It would then be appropriate to transmit a real-time traffic burst at level-1, and a TCP traffic burst at a lower level, say level-2. The burst header control packet would contain three pieces of information: the level in the hierarchy at which the burst will be transmitted, the start slot, and the burst length. A core node receiving this control packet would first deduce the outgoing link for the bursts, and then determine where the slot lies in its hierarchy corresponding to that output link. There are three possible outcomes:

1) A frame does not exist at the requested level in the hierarchy: For example, say Fig. 1 denotes the current hierarchy at the output link of interest at the core node, and say the new burst is arriving at level-2 starting in slot 8234. The 4-th slot in the level-1 frame does not have a level-2 frame under it, so there are two options: either create a new level-2 frame under this slot (if the slot is unoccupied), or use a delay line to delay the burst slices by one slot, moving it to the 5-th slot in the level-1 frame, which already has a level-2 frame underneath, and in which the 3-rd slot may be used if available.
2) A frame exists at the requested level but the required slot is unavailable: Again using Fig. 1 as an example, a new burst arriving at level-2 starting in slot 8375 collides with scheduled burst B_2. The new burst could be delayed using fibre loops by 10 slots to move it to the 8-th slot in the same level-2 frame. Alternatively, the new burst could be delayed by 3 slots to move it to the other level-2 frame if it has its 7-th slot available.
3) A frame exists and the requested slot is available: In this case the burst is assigned the requested slot and passes through the switch in a cut-through manner without any delays.

Several aspects of the architecture, such as the strategy for allocating and releasing appropriate slots in the hierarchy (i.e. is the allocation effective from the moment the burst header control is received or only for the duration of the burst data, etc.) and compaction of the hierarchy when all slots in a frame have been released are not discussed due to lack of space.

It is important to note that the complexity of control plane operations does not depend on the burst length; much like OBS (and TSOBS), bursts are scheduled atomically (not slice-by-slice) by finding an appropriate free slot in the hierarchy for the entire burst. This preserves the control plane scalability of OBS.

C. Data Plane Operation

The data plane uses the hierarchy constructed by the control plane for each output link. A counter is maintained for each frame in the hierarchy, corresponding to the slot last served in that frame. Each time-slot, the counter for the level-1 frame is incremented by one, and the corresponding slot entry checked. If it is a leaf entry containing a burst, the optical crossbar is configured so that the input line corresponding to that burst is switched to the output link under consideration. If on the other hand the slot entry points to a lower level frame, the counter for the lower-level frame is incremented, and the process recurses. Note that the optical delay lines are also scheduled by this process by treating them as output ports on the optical crossbar (e.g. in a "shared memory" architecture [4] where all fibre delay lines are connected to the central crossbar).

The complexity of the data plane operation per time-slot at most equals the number of levels in the frame hierarchy, which can be capped at a small constant. This preserves the scalability of OBS to high data plane rates.

III. CONCLUSIONS

In this paper we have presented an architecture for hierarchical time-sliced optical burst switching (HiTSOBS). HiTSOBS preserves the data and control plane scalability of OBS, while introducing a flexible frame hierarchy that allows different traffic classes to operate at different loss-delay trade-off points. An extended version of this paper will discuss the HiTSOBS architecture in greater detail, and also evaluate its performance via analysis and simulation.

REFERENCES

[1] C. Qiao and M. Yoo, "Optical Burst Switching (OBS) – A New Paradigm for an Optical Internet", *J. High Speed Networks*, 8(1):69–84, 1999.

[2] J. Ramamirtham and J. Turner, "Time Sliced Optical Burst Switching", *Proc. IEEE INFOCOM*, San Francisco, CA, Mar 2003.

[3] C. R. Kalmanek, H. Kanakia and S. Keshav, "Rate Controlled Servers for Very High-Speed Networks", *Proc. IEEE GLOBECOM*, San Diego, CA, Dec 1990.

[4] S. Yao, S. Dixit and B. Mukherjee, "Advances in Photonic Packet Switching: An Overview", *IEEE Comm. Mag.*, 38(2):84–94, Feb 2000.