Variable time-period optical switching: a novel OBS implementation

Jinsong Zhang (张劲松), Mingcui Cao (曹明翠), Fengguang Luo (罗风光), and Tao Chen (陈 涛)

State Key Lab of Laser Technology, Huazhong University of Science and Technology, Wuhan 430074

Received April 3, 2003

In this paper, we proposed a novel optical switching method based on optical burst switching (OBS), we call it variable time-period optical switching (VTPOS). It can both support circuit services and other immerged packet services. It has better usability of bandwidth, shorter offset and latency time than others of unidirectional transport signaling mechanisms for OBS. It supports deflection switching for improve blocking performance without the need of schedule buffer. It introduces a time pointer and phase indicator that made synchronous more precisely and requires less guard time, it also classifies the different services classes with a relative QoS model.

OCIS codes: 060.0060, 060.4250.

A recently alternative networking technology – optical burst switching (OBS)[1] is proposed. OBS representing a balance between circuit switching and packet switching indicates that an optical network solution delivers the benefits of optical packet switching (OPS) while avoiding the optical buffer memory and other hurdles.

In general, travelling on a separate control wavelength/channel, a burst control packet (BCP) is an offset time ahead of its data burst. The offset time is required because the BCP brings processing delay at each switching node. While the relative data burst does not need, the offset time is set to be the least.

The size of data burst (hundreds kilobits, even megabits) is usually far more than the size of IP packet. There is an idle guard time between two neighboring data bursts; it can eliminate the affection of imprecise synchronization clock. OBS scheme may be based on either two-way or one-way reservation protocol, but mainly based on one-way reservation protocol, in which a data burst follows a corresponding control packet without waiting for an acknowledgment, so it shortens the setup time and latency. For example, tell-and-go (TAG), just-enough-time (JET)[1] and just-in-time (JIT)[2], are all based on one-way reservation protocol[3].

Variable time-period optical switching (VTPOS) is a feasible OBS scheme based on current technologies. The basic scheme is as follows. A group of length-fixed timeslots (BCPs) constitute a burst control channel (BCP channel), multiple BCP channels, which are interlaced into an adapted payload structure in synchronization mode (SONET/SDH), constitute a BCP path (shown in Fig. 1). So a BCP path can provide BCPs for multiple data burst channels, which are carried on different data wavelengths. BCP channels will maintain their respective burst channels. Let T_{p} be the period of BCP path frame, and T_{b} be a basic time-period of a data burst, when $T_{\mathrm{b}} \geq T_{\mathrm{p}}$, each data burst on a data burst channel can be maintained by one BCP at least.

A synchronous BCP channel provides timeslot labels to mark BCPs sticking to their data bursts, and thus a BCP cannot link a wrong data burst even some errors happened. The synchronization clock signal on the BCP channel can be transmitted into all switching nodes in the network, so it can cut down the impairment of physical performance due to timing error. As BCP is the timeslot packet switching mode in the synchronous frame, the switching performance is better than that of a data burst; furthermore, a lost BCP only brings one basic time-period of resource bandwidth out of control.

The data burst scheme in VTPOS is a variable time-period. All data bursts in VTPOS will be an integral multiple of a basic time-period of burst (i.e. $n \times T_b$, $n = 1, 2, \cdots$). So, VTPOS can support both a fixed time-period data burst and a variable time-periods data burst, and multi-services can be transported directly in the VTPOS network with a less modification.

Due to the time delay of BCP processed in switching nodes, in Refs. [1] and [2], either JIT or JET, will keep enough offset time between BCP and its data burst at the source node, so as to ensure all intermediate switching nodes having time to process the BCP and implement switching. If $D(i)$ is the delay of processing time of BCP with traffic class i in each node, and H is the max number of switching nodes along the switching route, $Q(i)$ is the offset time at the source node, then $Q(i) \geq D(i) \times H$. Thus, it reduces the efficiency of utilization of the network bandwidth and increases the latency, additionally; either of them does not support the deflection switching mode, otherwise, processing in additional switching nodes makes a data burst run ahead of its BCP. In VTPOS, instead of a long offset time between a BCP and its data burst, a reasonable minimum offset time is set in any switching node along with the switching path. There

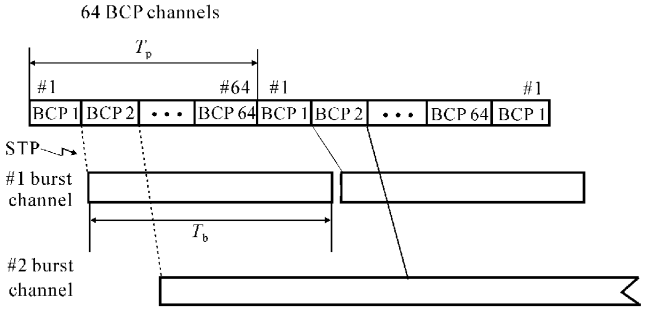


Fig. 1. A BCP path with synchronous frame structure.

is a synchronous time pointer (STP) in a BCP (see Fig. 2) which indicates the offset time (T_f) precisely. For example, at ingress of a switching node k, the previous STP is $T_\mathrm{f}(k-1)$, at first, the data burst will delay a fixed processing time-period $D(k)$, the BCP actual processed time is $T_\mathrm{p}(k)$, STP can be adjusted to the actual time $T_\mathrm{f}(k) = T_\mathrm{f}(k-1) + D(k) - T_\mathrm{p}(k)$ at egress of each switching node. When the time interval is longer than a basic time-period of data burst (i.e. $T_\mathrm{f}(k) > T_\mathrm{b}$), the BCP can slip backward the following idle BCP slot of the same BCP channel. Simultaneously adjusting a T_p of variation (the process is shown in Fig. 3), so that the STP $T'_\mathrm{f}(k) = T_\mathrm{f}(k) - T_\mathrm{p}$, will be kept a reasonable minimum range.

By setting a fixed fiber delay line on the path of data burst at the input of a switching node in VTPOS, the switching processing time interval will be distributed equably in all switching nodes that the burst will pass by. The advantages of the processing scheme of the offset time are that, it will reduce the offset time in each switching node, and it has the convenience for scheduling, processing flexibly, and resource distribution in any switching node. On the other side, while a data burst may be blocked at one switching node without optical buffer, the reasonable minimum offset time between the BCP and its data burst makes it possible to support the deflection route switching in VTPOS, and thus to reduce the blocking probability of the burst.

Unlike other OBSs, VTPOS can support two switching modes, the channel automatic delay released (ADR) mode and the channel claimed interruption (CI) mode (see Fig. 4).

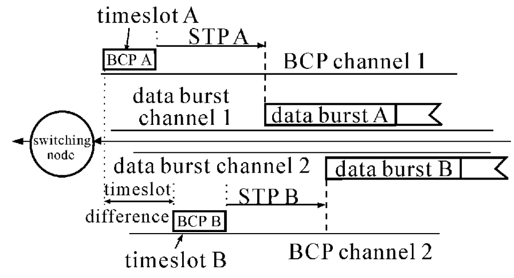
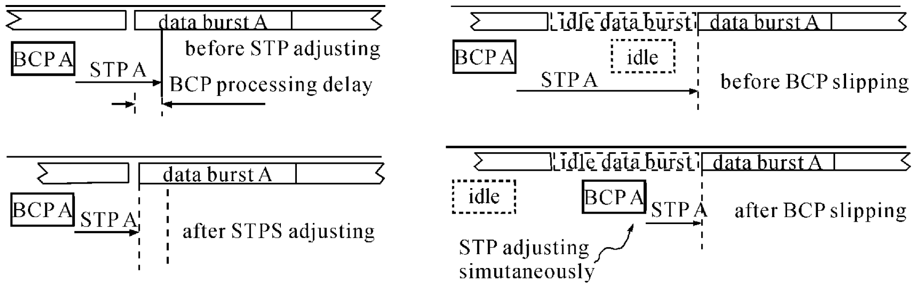
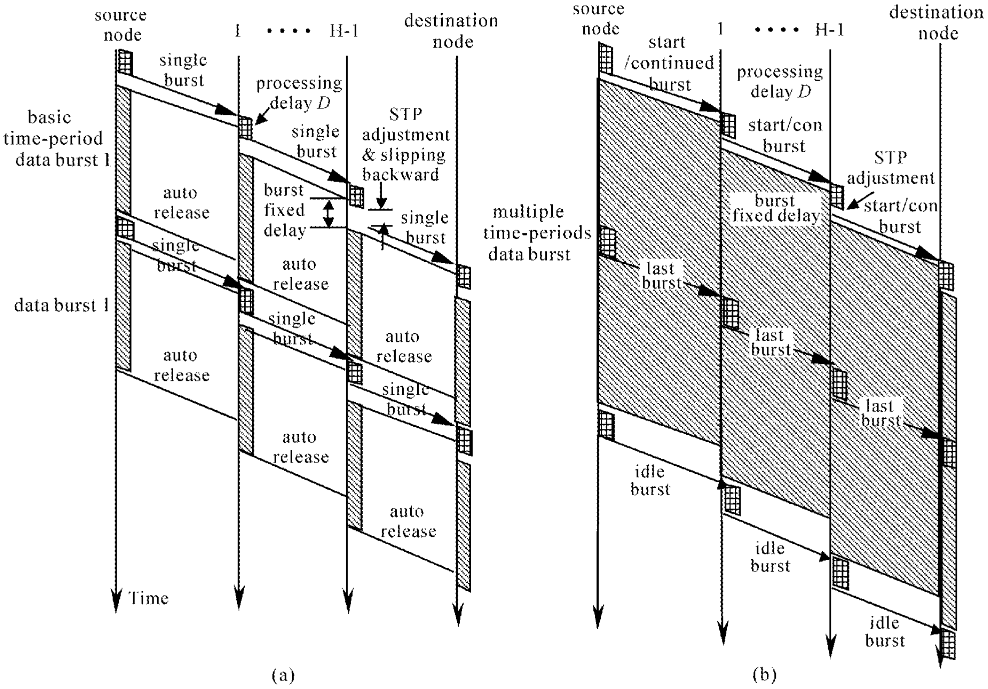


Fig. 2. STP in BCP.



(a) (b)

Fig. 3. Process of STP adjusting and BCP slipping backward: (a) STP adjusting without BCP slipping; (b) STP adjusting with BCP slipping backward.



(a) (b)

Fig. 4. Signaling protocols of VTPOS: (a) signaling of service class A (under ADR mode); (b) signaling of service class B (under CI mode).

Under the channel ADR mode, actually, it supports a single basic time-period burst switching. After a burst passing by, the switching node reclaims the reserved bandwidth resource automatically, the data burst channel will be interrupted and bandwidth resource will be drawn back; the switching node begins to process the next BCP and its data burst. Under the channel CI switching mode, when the data burst of some service passes through the switching node, the switching resource and state are reserved for the same service burst until a confirmed releasing BCP comes. Meanwhile, long or short bursts time can be available, furthermore, high-speed circuit services can be transmitted directly, and therefore, it can improve the efficiency of bandwidth utilization.

Furthermore, VTPOS provides a different service of classes indicator for bursts in a relative quality model. Intermediates can distinguish up to 16 levels of service classes. A high-level of data bursts is prior to low-level one to pass through a blocking node. In order to ensure the burst of high priority being on a lower block probability, delay-switching windows will be introduced into the implement of the switching schedule in nodes.

For an implementation paradigm of VTPOS, the control wavelength (carrying logical BCP channels) is λ_0, it works at STM-4 (STS-12) in SDH (SONET) frame[5], instead of ATM channels carrying BCP channels (out of band), and the data wavelengths (carrying logical data burst channels) are λ_i ($i = 1 - 64$, the bit rate is 10 Gb/s), all BCPs in one control wavelength will cope with up to all data bursts in 64 data wavelengths. But these channels of BCP are different from the channels of which are used to transport the circuit payload. The average time-period of BCP is about 3.472 μs, which is less than the frame time period of VCs (from 125 to 500 μs).

Figure 5 shows BCP path frame structure based on STM-4. A four-byte BCP that will bring a shorter processing time at switching nodes is shown in the figure. Otherwise, a mount of reserved overheads (such as ROH, MSOH, VC-4-4c POH) will provide system OAM and common resource signaling control. The time-period of BCP is only 1/36 of VC-4 time-period, and each path frame contains 65 BCP channels, from Ch0 to Ch64. Ch0 is reserved for transmission of system control protocol, Ch1 to Ch64 are used as BCP channels, which are associated with corresponding data wavelengths. When the time-period of each BCP channel is shorter than time length of the basic data burst, 4 bytes BCP in a BCP channel can serve all data bursts in the data burst channel. The main field functions in a BCP are explained as follows.

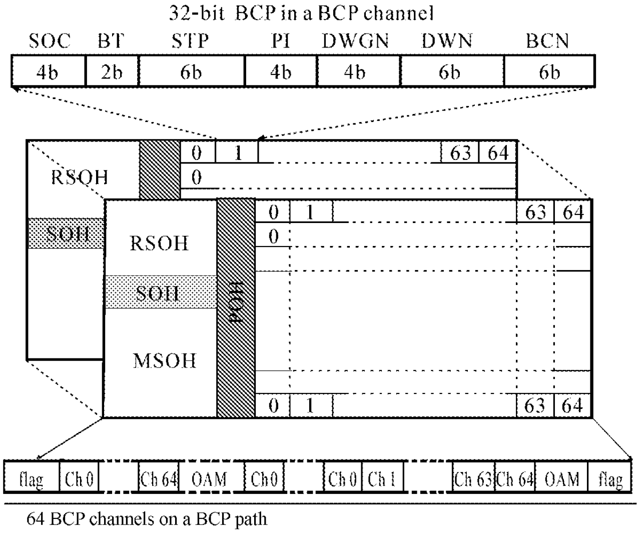


Fig. 5. BCP path structure based on STM-4 (OC-12).

Service of classes (SOC, 4 bits): it is a level indicator of relative quality model. It can distinguish with 16 levels of service class. High-level of data bursts are prior to low-level ones to pass through a blocking node. All '1's (1111), the highest class with CI mode, directly supports circuit service; 1110, the second highest class with ADR mode; from 1101 to 0001 indicate other classes and modes (it will be researched in the future); all '0's (0000), the lowest class burst (an available burst), while some nodes meet traffic blocking, the class burst will be dropped firstly. On the other side, the integrated number is composed by SOC and BCN, which marked a burst connection number in a respective data burst channel. SOC comprises the high part of the number.

BT (2 bits): it is reserved for indicating burst type, including idle burst (00), start & continue burst (01), single burst (11), and last burst (01).

Synchronous time pointer (STP, 6 bits): the time interval between BCP with the beginning of its relative data burst. STP can be modified at egress of each switching node according to the real time changing. Each step adjusting of STP represents a sixteenth of BCP time-period changing.

Burst phase indicator (PI, 4 bits): it provides more precisely interval time indicator between the BCP and its data burst.

Burst connection number (BCN, 6 bits): it indicates a logical burst connection of the data burst channel. DWN is local physical label of a data wavelength carrying a data burst channel. DWGN and DWN are local physical label of a data wavelength group.

Aiming at introducing a basic COS in the VTPOS network, delay switching windows schedule is proposed. In the VTPOS network, the BCP will arrive several basic time periods ahead of its data burst (sometimes using some FDL to adjust time interval between BCP and its data burst) at a switching node, according to the different classes and length of data bursts, the data burst will be on the different priorities and different length of switching windows. Between the BCP and its data burst, thus, when the blocking happens, the shorter and lower burst will be dropped firstly. Under the mode, there is no switching buffer required in optical domain.

For example, Fig. 6 illustrates the data bursts with different classes passing by a switching node. There are three classes of services with different priorities and lengths of delay switching windows, bursts A, B and C, lengths of delay switching windows are DSW1, DSW2, and DSW3 respectively. Priorities are higher gradually, at given time, they switched to the same destination path at the node, arriving times of three services are described in the figure. From the figure, we know that burst A arrives firstly; it will be transferred to the pre-output

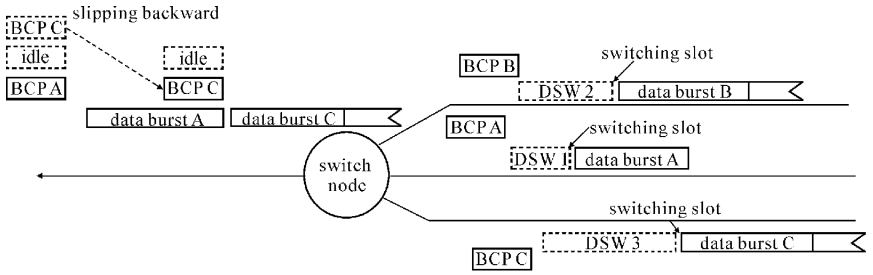


Fig. 6. Delay switching window.

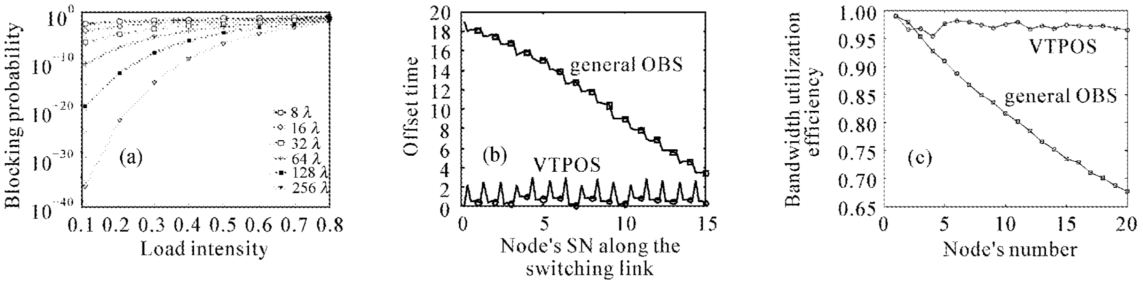


Fig. 7. Performance simulations of VTPOS.

queue. Subsequence, burst B arrives, because the priority is higher than burst A, and the two bursts will be contention, so burst B is transferred to the pre-output queue, and burst A is fetched and placed to the waiting switching queue. Later, burst C arrives. it will be compared with those bursts both in the pre-output queue and the waiting switching queue on timing relations. As a result, bursts A and C are transferred to the pre-output queue, and burst B is transferred to the waiting switching queue inversely. As soon as the switching slot comes, burst B is dropped; and bursts A and C pass the switching node.

Now we will validate our analytical results of some parameter performance by simulation.

Firstly, we calculate the blocking probability under different numbers of wavelengths in a switching node without optical buffer in VTPOS. It is assumed that the burst arrival processes at each input port are Poisson arrival, the time-period of burst is exponential and general distribution, and each input link is given the same traffic load intensify ρ, destination addresses of the bursts are uniformly distributed over all of the outgoing links of the optical switching node. So the traffic of simulation follow an $M|M|n|n|m$ traffic load model system[6], n is shared wavelength number, it is a half of m.

Figure 7(a) plots the blocking probability versus traffic load intensity for the different numbers of wavelengths m. From the result, we know when the data burst wavelength is 64, the load is 0.5, and we can get an acceptable blocking probability of about 10^{-4} level. Figure 7(b) shows the offset time between a BCP and its burst for general OBS and VTPOS along the switching path, respectively. The number of switching nodes k varies from 2 to 20. We know that the offset time for general OBS varied greatly from the source node to the destination node, due to keeping reserved processing time enough. But the offset time for VTPOS is short and with little variation, because it can be changed by fixing FDL and adjusting the BCP in each switching node. Figure 7(c) shows that different offset time will get different optical bandwidth utilization efficiency under delay reservation and delay switching windows methods.

In this paper, we proposed a new optical switching method based on OBS – VTPOS. We presented briefly the architecture, signaling, design, and implementation for VTPOS. It can both support circuit services and other immerged packet services. It is an optical burst time switching that is independent of the signal types and the service's bit rate.

At the aspect of usability of bandwidth, it will be better than general OBS protocols of unidirectional transport signaling mechanisms for OBS. No schedule buffer is required. In VTPOS, BCP channel is based on SDH structure, so it can reduce timing error between the data burst and its BCP. The synchronous time pointer is introduced to get more precise synchronization and less guard times, as well as less offset time and latency. These increase the optical bandwidth utilization efficiency. It also classifies the different services classes with the relative QoS model, improves the performance of higher classes of services.

F. Luo is the author to whom the correspondence should be addressed, his e-mail address is fgluo@mail.hust.edu.cn.

References

1. M. Yoo, C. Qiao, and S. Dixit, IEEE Commun. Mag. **39** (2), 98 (2001).
2. J. Y. Wei and R. I. McFarland, Jr., J. Lightwave Technol. **18,** 2019 (2000).
3. M. Yoo, M. Jeong, and C. Qiao, Proc. SPIE **3230,** 79 (1997).
4. M. A. Marsan, A. Bianco, P. Giaccone, E. Leonardi, and F. Neri, IEEE/ACM Trans. on Networking **10,** 666 (2002).
5. ITU-T Recommendation G.707, 3 (1996).
6. H. L. Vu and M. Zukerman, IEEE Commun. Lett. **6,** 214 (2002).