A novel and simple beforehand bandwidth reservation (BBR) MAC protocol for OBS metro ring networks

Wen-Ping Chen, Wen-Fong Wang and Wen-Shyang Hwang

Abstract. This paper proposes a novel and simple MAC protocol called beforehand bandwidth reservation (BBR) to reserve the empty slots in the next big-slot cycle for OBS ring networks by control channel. The node architecture uses one tunable transmitter and one fixed receiver to add or drop data channels and a fixed transmitter/receiver pair to transmit and receive on the control channel. Each node possesses a dedicated channel to receive data, so it inherently occupies a priority position. Fortunately, a well-known packet scheduling approach, time-division multiplexing (TDM), can overcome this problem. In addition, the length of the big-slot cycle is studied with a view to investigating performance divergence. Finally, a multi-token protocol using the (FTW–FRW) node architecture is compared with the BBR protocol. According to simulation results, the BBR scheme using the TDM approach achieves the best bandwidth utilization (more than 95%), while the multi-token protocol achieves the worst performance based on packet distribution in the MCI backbone.

Keywords: Beforehand bandwidth reservation, packet scheduling, TDM, OBS, ring networks

1. Introduction

In the past decade, Internet applications such as electronic commerce, multimedia file sharing, voice-over-Internet protocol (VoIP), and storage area networks (SANs) have led to an enormous increase in bandwidth requirements. The need for a transmission medium with the bandwidth capabilities for handling such a vast amount of information is critical. The need to provide the bandwidth necessary to fulfill this ever-increasing demand and accommodate rapid advances in optical technologies, especially wavelength-division multiplexing (WDM) technology, has been driving the evolution of Internet architecture. In the next-generation optical Internet, optical technology will play a more dominant role, not only in transmission, but also in switching.

To date, research activities have been focused on two optical switching technologies: optical burst switching (OBS) [1–7] and optical packet switching (OPS) [8–14]. However, OPS technologies are still immature because of technical challenges in developing optical RAM, optical processing, and optical switching. OBS is proposed as an improvement to this immature switching technology by configuring the optical switching elements in advance; it adopts electronic processing for the control signals that carry all the required information for the switched optical bursts [1]. Furthermore, the choice of signaling protocol is one of the most important issues for collision-free packet transmission and bandwidth reservation. Just-Enough-Time (JET) [1,2] and Just-In-Time (JIT) [3] are two feasible signaling protocols which have been investigated in OBS research. In mesh networks, when a burst packet
is transmitted, it is difficult to know how much offset time to allow to avoid collisions between the source node and
the destination node. Therefore, the JIT protocol is difficult to implement, so most studies adopt the JET protocol to
implement OBS mesh networks. However, in ring networks, the JIT protocol is currently used, because the process
time and delay time from source node to destination node can be calculated.

In the past years, the OBS ring networks are proposed in [4–7]. However, in [4–6], the authors adopted a
token-passing scheme to implement collision-free transmission, but in such systems, the end-to-end delay is highly
redundant because the token-holding packet stays in the token queue for a long time. In [7], the authors proposed
two schemes: the dynamic OBS-JET ring and the dynamic WR-OBS ring. However, the WR-OBS ring suffers
from the fatal flaw that all systems would experience failure whenever the central node malfunctions. In the OBS-
JET ring, it is difficult to avoid packet collision. As a result, to cope with the requirements of next-generation metro
networks, this paper proposes a novel and simple MAC protocol, beforehand bandwidth reservation (BBR), for an
OBS ring network. The protocol uses a control channel to reserve the empty slots in the next big-slot cycle to
avoid packet collision. It can support variable-length packets without complex fiber delay line buffers, complexity
algorithms, or optical-electronic translation from source node to destination node. In addition, to simplify and
reduce the cost of node architecture, the tunable transmitter–fixed receiver (TT–FR) architecture has been adopted.
However, an unfairness problem intrinsic to the concept of position priority exists. Fortunately, this space problem
can be addressed by including the concept of quotas from the statistical-TDM scheme, but the length of the big-
slot cycle has a major influence on the system, because its size as implemented is too small to add the overhead
of the tunable laser which adjusts the working wavelength. On the other hand, if the big-slot cycle is made longer,
this will create a long end-to-end transmission delay. Therefore, optimizing the length of the big-slot cycle is a
very important issue in the BBR protocol. Here, three kinds of different packet scheduling algorithms: starvation,
synchronous-TDM, and statistical-TDM schemes have been proposed to investigate this issue based on packet
distribution in an MCI backbone with OC-3 links [15].

In the rest of this paper, the WDM metro ring network architecture for the proposed protocol is presented in
Section 2. The BBR MAC protocol involving three kinds of packet scheduling for supporting OBS metro ring net-
works is presented in Section 3. Section 4 describes the system model and simulation assumptions. The simulation
results and a discussion of a broad range of related issues are presented in Section 5. Finally, a few remarks are
given in the conclusion.

2. System architecture

Today, using current technology to realize purely optical packet processing is difficult and inefficient because of
the lack of practical solutions for optical memory and high-speed optical switching. However, advances in tunable
lasers, filters, optical circulators, and fixed Bragg gratings (FBG) have made rapidly reconfigurable lightpaths a
reality. Therefore, for supporting bursty and rapidly changing data traffic, this paper proposes a network archi-
tecture that supports variable-length data burst transmission via dynamic bandwidth allocation of slot time on a
WDM ring network, as shown in Fig. 1. The network architecture presented here is based on a single unidirec-
tional fiber ring which connects \( N \) nodes. The optical fibers between nodes are each composed of \( W + 1 \) channels
\( (\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_W, \lambda_c) \). Here, each access point (AP) in the LightRing [13] network is equipped with one tunable
transmitter and fixed receiver (TT–FR) for data channels and a fixed-transmitter and fixed-receiver pair (FT–FR)
for the control channel in an OBS ring network.

Each node has one dedicated data channel to receive data from other nodes. The number of nodes, \( N \), in this
network is equal to the number of channels, \( W \). Furthermore, the ring is assumed to be of wide scope, covering a
major metropolitan area (i.e., its circumference is about 100 km). The whole system is referred to as a WDM metro
ring. In the network, access-point (AP) nodes are used to connect a LAN to the MAN, while a “point of presence”
(PoP) node connects the MAN to a WAN. The AP may exchange burst packets with an attached LAN, a gigabit
Ethernet (GbE) network, a passive optical network (PON), or a wireless LAN. Burst packets travel along the ring
as optical signals, without any need for electro-optic conversion at the intermediate nodes.
Fig. 1. BBR-OBS architecture of a WDM metro ring network.

Packets to be sent from a LAN are first inserted into one of \((N - 1)\) first-in first-out (FIFO) transmission queues, according to their destination address. Each node has a dedicated wavelength to receive the burst packet from another node. The node architecture can be treated as that of an optical add/drop multiplexer (OADM) node. In its function as a drop model, each AP installs one optical circulator and two FBGs to reflect the drop burst packet and the control packet. In its function as an add model, each AP node adds the local traffic to the upstream home channel using a combiner. It is important to note that the home wavelength can be shared among several nodes during both reception and transmission. Here, in order to avoid collisions among burst packets transmitted in the home channel, a beforehand bandwidth reservation (BBR) MAC protocol based on a slot-time scheme is proposed. The protocol uses the control channel to carry the status of empty or full slots during the next big-slot cycle which consists of numerous slots. Each node complies with the message to reserve the empty slots and transmit a burst packet associated with the particular destination at the next big-slot cycle. A fiber delay line (FDL) is installed to delay the incoming data for a fixed time to recognize the status (empty or full) of all slots in all data channels by a control message in the next big-slot cycle, and to allow necessary processing time for the MAC protocol. Afterwards, a new control message will be transmitted over the fiber to renew the original control message. Here, the wavelength time \((T_W)\) [19] required by the tunable laser to adjust the working wavelength to the new wavelength is also considered.
in practice. The advantage of this node architecture is that it renders unnecessary the use of immature optical switch arrays or complex optical delay line arrays to complete data add/drop operations. It needs only a small FDL to satisfy the requirements of the MAC protocol and the ability to use a home channel to provide a fractional wavelength capacity to each ring node, reducing maintenance costs due to the use of passive devices.

3. MAC protocol

This article proposes a novel and simple MAC protocol. It belongs to the class of collision-free access protocols and is called beforehand bandwidth reservation (BBR). It is based on a time-slot scheme and is intended to support variable-length burst packets over WDM OBS ring networks. In addition, to use empty slots fully and fairly, three kinds of packet scheduling (starvation, synchronous-TDM, and statistical-TDM) are also investigated. Finally, the frame format of the BBR protocol will be described in detail.

3.1. BBR MAC protocol

In metro ring networks, performance depends mainly on the sharing of optical resources among the various competing access nodes. Consequently, both collision-free transmission and fair sharing of the common bandwidth among ring nodes must be ensured. Here, a MAC protocol, beforehand bandwidth reservation (BBR), is proposed. The protocol uses one control channel to process the bandwidth reservation in the next big-slot cycle which consists of numerous slots, with the intent of avoiding collisions. On the circumference of a single ring, many big-slot cycles continuously circulate around the ring, as shown in Fig. 2. According to the generic slotted-ring access mechanism, each big-slot cycle in the control channel records the status of slots (empty or full) at the next cycle on all data channels ($\lambda_1, \lambda_2, \ldots, \lambda_W$) using subcarrier multiplexing label technology [18]. The corresponding control packet records two states (“0 = empty” or “1 = occupied”), thus capturing the status of each data channel. If the AP$_i$ wants to transmit its data to node AP$_j$, it must first identify how many empty slots in the data channel can be captured in the big-slot cycle which now occupies the control channel. If $P$ slots are empty, and it is assumed that $T_W$ is wasting $q$ slots, then $P - q$ slots can be used by AP$_i$. Afterwards, the maximum number of “0” bits in $P - q$

![Fig. 2. Data and control channel messages in AP$_1$.](image-url)
Fig. 3. Data and control channel messages in AP1 using starvation TDM scheme.

can be changed to “1” in the corresponding channel field of the control message, and an assembled burst packet will be transmitted in the corresponding data channel at the next big-slot cycle.

However, in this protocol, the allocation of the empty slots by each node during one big-slot cycle can be achieved by any of three well-known packet scheduling algorithms: starvation, synchronous-TDM, and statistical-TDM (time-division multiplexing). Here, for purposes of a simple description, we assume that $N = W = 4$, the $T_W$ distance is equal to one slot, and $S = 21$ slots in one big-slot cycle. The data and control channel message in AP1 according to starvation TDM scheduling is shown in Fig. 3. Obviously, AP1 has a high priority to capture the empty slots on the fourth data channel because of its high-priority position. As a result, it is very easy to transmit the data to AP4. On the other hand, it is difficult for AP1 to capture the empty slots on the second data channel because of its low priority, particularly under heavy load. This approach apparently produces an unfair transmission allocation. Fortunately, the synchronous-TDM scheme can overcome this problem. Figure 5 shows that each node is assigned the same quota value to capture the empty slots in every data channel during one big-slot cycle. The quota value is:

$$Quota = \frac{S - q(W - 1)}{W - 1} = \frac{S}{W - 1} - q.$$  \hspace{1cm} (1)

The total number of big-slot cycles around the ring and the maximum bandwidth utilization for each data channel are:

$$M = \frac{L}{C} \times \frac{R}{S} \times l,$$  \hspace{1cm} (2)

$$\eta_{\text{max}} = \frac{(W - 1)Quota}{S} = \frac{W - 1}{S} \times \left( \frac{S}{W - 1} - q \right) = 1 - \frac{q}{S},$$  \hspace{1cm} (3)

where: $l$ is the bit length of the slot, $L$ is the length of the ring network, $\eta_{\text{max}}$ is the maximum utilization, $W$ is the number of data channels, $R$ is the bit rate in each data channel, $C$ is the velocity of light in the fiber, and $S$ is the total slots in each data channel in the big-slot cycle.
However, if the value of the above formula (2) is not an integer, the remaining slots are added into the last big-slot cycle, enabling the BBR protocol to survive on any length of ring network.

Although the synchronous-TDM scheme can maintain fairness, it can still be improved with respect to throughput and end-to-end delay. Figure 4 shows an occurrence of waste which arises because the fourteenth slot of the \((n - 1)\)st big-slot cycle on \(\lambda_1\) and the sixth and seventh slots of the \(n\)th big-slot cycle on \(\lambda_2\) have not been used by AP3, even if AP4 has enough time to tune the wavelength. Therefore, the statistical-TDM scheme is proposed to address this problem. This scheme not only increases the utilization rate of the empty slots, but also reduces the end-to-end delay under light load. Figure 5 shows an example in which the number of slots captured by AP4

![Fig. 4. Data and control channel messages in AP1 using synchronous-TDM scheme.](image1)

![Fig. 5. Data and control channel messages in AP1 using statistical-TDM scheme.](image2)
are increased for the \((n - 1)\)st and \(n\)th big-slot cycles. The synchronous-TDM algorithm can be written as follows:

**Statistic-TDM Algorithm**

**Procedure** Synchronous-TDM

let \(\text{quota}(j) = \frac{S}{W - 1}\)

for \(i = 1\) to \(W\)

for \(j = 1\) to \((W - 1)\)

if the remains of \(\{\text{quota}(i, j) - q\}\) \(\geq\) the remains of \(\text{quota}(i - 1, j)\) and the \((j + 1)\) transmission queue has packets wait for transmitting

then the extra \(\{\text{quota}(i, j) - \text{quota}(i - 1, j) - q\}\) empty slots can be used by \(\text{AP}_i\)

end if

\(j = j + 1\)

loop

\(i = i + 1\)

loop

In the BBR protocol, there is an important factor \(S\) in packet scheduling influences the system performance; the smaller \(S\) is the more ratio of \(T_W\) in a big-slot cycle. The result will be a reduction in bandwidth utilization, even though the algorithm does not have enough time to recognize the status of empty and full slots in each big-slot cycle and to execute the MAC packet scheduling process. On the contrary, a very large value of \(S\) will result in a long end-to-end delay. Therefore, the appropriate design of \(S\) is a worthy topic of investigation and is simulated and discussed in Section 5 of this paper.

To support the BBR protocol, we have designed the frame format shown in Fig. 6 to solve the addressing problem. It consists of a start delimiter (SD) to label the channel conveying the data frame, source address (SA), and a frame length (FL) field indicating the length of the packet. Finally, a cyclic redundancy check (CRC) field is used to detect and correct errors.

### 4. System model and assumptions

In order to understand the performance of the BBR-ring network system, this section will analyze the average transfer delay and maximum throughput in the multi-ring cases under the following assumptions.

(a) The number of WDM channels (or logical rings) is \(W\), and each data channel connects one designated node.

(b) The nodes are equidistant, and the number of packets being transmitted and received by each node is balanced. The mean transmission distance is thus \(N/2\) nodes.

(c) Packets arrive at node \(i\) according to an independent, identically distributed \((i.i.d.)\) Poisson process at the rate \(\lambda_i\). Hence, the aggregate arrival rate of the network is \(\lambda = \sum_{i=0}^{N-1} \lambda_i\).

(d) The distribution of arriving packets at node \(i\) destined for node \(i \oplus j\) is also determined by a Poisson process with the rate of \(\lambda_{i;i \oplus j}\), as shown in Fig. 7. The \(\oplus\) symbol indicates addition modulo \(N\); \(\lambda_i = \sum_{j=1}^{N-1} \lambda_{i;i \oplus j}\).
(e) Each node generates packets at the same average rate, and each sends an equal amount of traffic to all destinations (uniform and symmetric traffic):

\[ \lambda_i = \frac{\lambda}{N}, \quad \lambda_{i, i+j} = \frac{\lambda_i}{N-1} = \frac{\lambda}{N(N-1)}, \quad (4) \]

where \( \lambda_{i, i} = 0, \) for \( 0 \leq i \leq N - 1, \) \( 1 \leq j \leq N - 1. \)

(f) The lengths of the packets are randomly determined by independent, identical, geometrically distributed random variables (denoted by r.v., with \( M \) bits), with mean \( E[M] \) and the following probability mass function [16]:

\[ P_r(M = k) = \beta \times (1 - \beta)^k, \quad k = 0, 1, 2, \ldots, \]

where \( \beta = 1/(1 + E[M]). \)

(g) The length of any transmission queue is infinite.

Under these assumptions, queuing and transmission delays can be modeled using an M/G/1 queue. We also compute the following performance metrics:

(a) The throughput per node \( (P) \) is the total number of bits sent from a node divided by the simulation time.

(b) Mean packet transmission delay \( D = \) mean queuing delay + mean propagation delay + mean transmission time.

5. Simulation results and discussion

The performance of a BBR ring network is simulated in SIMSCRIPT II, using IP packets with a size distribution that matches the trace observed in an MCI backbone with OC-3 links (Fig. 8). The mean packet size in this distribution can be calculated as 353.8 bytes. Here, it is important to note that all simulations were run until the network reached steady state. In general, each point in the plots below corresponds to approximately ten million time units of simulation. In addition, each experimental value is calculated by the variance reduction technique using 40 replicated simulations with different random seeds. The full simulation parameters for OBS metro ring networks using the multi-token and BBR MAC protocols are given in Table 1, including, for example, the wavelength of the GCSR laser [8] which is adjusted to 4–10 ns. Assuming a channel speed of 10 Gb/s, \( T_W \) consumes a bandwidth of only 5–12.5 bytes. Fortunately, the length of one slot is enough to fulfill the request.

5.1. Performance evaluations of the multi-token and BBR-ring protocols

Figure 9 shows the mean transmission delay versus the imposed load under the multi-token ring protocol for various lengths of ring. It is clear from Fig. 9 that the mean throughput of the network depends on ring length.
because in a longer network, the token packet stays in the fiber for a longer time. However, the performance of the multi-token protocol is very poor on ring networks. The mean throughput of each node is less than 2 Gb/s if $L$ is large (100 km) under $N = W = 4$ and transmission delay = 10 ms. However, if $L = 10$ km, the throughput of each node reaches almost 10 Gb/s under the same conditions. Therefore, the multi-token protocol is unsuitable to support a metropolitan area network, but is suitable only for small-area networks such as local area networks.

Figures 10–12 illustrate the mean transmission delay versus the imposed load under the BBR protocol with the three packet scheduling algorithms for various numbers of slots in each big-slot cycle. It is obviously that the relative throughput delay of the network depends on the size of $S$. Considering both throughput and delay, a value of $S$ between 500 and 600 leads to better performance with $W = N = 4$. For example, the throughput of each node with the starvation-TDM algorithm is 9.0 Gb/s, with synchronous-TDM a rate of 9.2 Gb/s can be attained, and the statistic-TDM can obtain the best performance, 9.5 Gb/s under $W = N = 4$, with transmission delay = 10.0 ms and $S = 500$. These packet scheduling algorithms all have better performance than that of the multi-token protocol. Figure 13 illustrates the throughput comparison under the BBR protocol of the three packet scheduling algorithms as well as the multi-token protocol. Figure 13 clearly shows that the BBR protocol with statistical-TDM is much better than the other two packet scheduling algorithms or the multi-token protocol.
Fig. 9. Mean transmission delay of multi-token protocol for various lengths of ring, as a function of the imposed load.

Fig. 10. Mean transmission delay in the BBR ring using starvation-TDM with $N = W = 4$, $L = 100$ km, and various values of $S$, as a function of the imposed load.

However, a high burst length can reduce the overhead ratio for transmission data, but it also adds assembly/disassembly time and increases the retransmission time when the burst experiences loss or damage between the source node and the destination node. Therefore, this paper also studies and compares the relationship of average burst length to imposed load under the BBR protocol using three packet scheduling algorithms and the multi-token protocol, as shown in Fig. 14. It is evident that the average burst length in the BBR ring with all three packet scheduling algorithms remains below 9 kbytes even under heavy load. However, with the multi-token protocol, the average burst length increases with greater imposed load because the TX queue must wait for a longer and longer time to capture the free token.

Figures 15 and 16 illustrate the mean transmission delay and mean burst length under the BBR protocol using statistical-TDM with $N = W = 4$ and various lengths of ring, as a function of imposed load. It is evident that the influence of performance and mean burst length remains constant with changes in ring length. Because the size of $S$ is the only main influencing factor under this protocol, the ring length barely affects the packet propagation delay.
Fig. 11. Mean transmission delay in a BBR ring using synchronous-TDM with $N = W = 4, L = 100$ km, and various values of $S$, as a function of the imposed load.

Fig. 12. Mean transmission delay in the BBR protocol using statistical-TDM with $N = W = 4, L = 100$ km, and various values of $S$, as a function of the imposed load.

From the above discussion, the BBR protocol with statistical-TDM has the best performance, with a throughput per node more than five times greater than the multi-token ring under the same conditions: $N = W = 4, L = 100$ km, and transmission delay = 10 ms. With regard to the bandwidth utilization of all the OBS ring networks, the proposed protocol exhibits excellent characteristics of high throughput and low delay for all optical communications.

6. Conclusions

A novel and simple MAC protocol, beforehand bandwidth reservation (BBR), based on a time-slot scheme is proposed for all optical OBS ring networks. The protocol can provide collision-free transmission, high utilization,
and support for variable-length packet transmission over WDM from LAN to MAN. In addition, the allocation of empty slots in a big-slot cycle to increase performance effectively and fairly under the BBR protocol with three kinds of packet scheduling algorithm has been investigated. Simultaneously, this paper has also investigated the performance of the multi-token protocol as a benchmark against which the new protocol can be compared. According to simulation results, the BBR protocol with statistical-TDM scheduling has better performance than the other packet scheduling algorithms and the multi-token ring network. Not surprisingly, it can be observed that the throughput characteristics of the network are related to the number of slots in each big-slot cycle. Simulations show that values of $S$ between 500 and 600 with $N = W = 4$ offer better performance with respect to both throughput and delay. Network performance can approach 9.5 Gb/s in the case of $S = 500$. What is more, the average burst length is smaller than with the multi-token protocol, in which the TX queue must often wait for a long time to capture the free token. The advantages are avoiding the redundant assembly/disassembly time and retransmission time when a burst packet experiences loss or damage between the source node and the destination.
node. With regards to bandwidth utilization in all the optical ring networks, the proposed protocol exhibits excellent characteristics of high throughput and short delay for all optical communications.

Acknowledgment

The authors would like to thank National Science Council (NSC) in Taiwan of R.O.C. for supporting this research; the project number is NSC 95-2221-E-151-029.

References


