A Novel Burst Assembly Algorithm for OBS Networks Based on Burst Size and	Thuật Toán Lập Lịch Mới Lạ cho Các Mạng OBS Dựa Trên Dự Đoán Kích

Assembly Time Prediction checked By using a hybrid threshold to promptly detect the alternation of traffic load, the size threshold can be adjusted adaptively by the step size. Time Prediction Equation is investigated and demonstrated to be effective to control the prediction success probability, While the first IP packet is coming into the assembly queue, both data burst assembly size and time are immediately predicted respectively according to the size threshold and Time Prediction Equation, in order to reduce the latency and avoid unsuccessful data burst prediction. An aggressive algorithm is proposed to deliver better performance improvement with latency reduction and bandwidth utilization. The simulation results show that this assembly algorithm can reduce the End-To-End delay significantly without any additional time overhead, and any bandwidth wastage due to unsuccessful prediction. 1 Introduction Optical Burst Switching (OBS) [1, 2] is positioned between Optical Circuit Switching (OCS) and Optical Packet Switching (OPS), and can be implemented by separating the control channel from data channel. Reservation of networks resources, such as wavelength converters or data channels, is performed at each intermediate core node after receiving and interpreting the burst control packet (BCP). When the data burst (DB) is coming, the node has configured its switch fabric well. For these reasons, OBS doesn't require the	Thước Burst và Thời Gian Lập Bằng cách sử dụng ngưỡng lai hóa để phát hiện nhanh sự thay đổi lưu lượng đường truyền, ngưỡng kích thước có thể được điều chỉnh thích ứng bằng kích thước bước. Phương trình dự đoán thời gian được khảo sát và chứng minh là hiệu quả trong việc điều khiển xác suất dự đoán thành công, Trong khi gói tin IP đầu tiên đến hàng đợi lập burst, cả kích thước lập và thời gian của burst dữ liệu được dự đoán ngay lập tức theo ngưỡng kích thước và Phương Trình Dự Đoán Thời Gian để giảm độ trễ và tránh dự đoán burst dữ liệu không thành công. Chúng tôi đề xuất thuật toán tích cực để cải thiện thêm hiệu suất đồng thời giảm độ trễ và tăng khả năng tận dụng băng thông. Kết quả mô phỏng chứng tỏ rằng thuật toán lập burst này có thể giảm độ trễ đầu cuối đáng kể mà không tăng thêm thời gian lãng phí, và bất kỳ sự lãng phí băng thông nào do dự đoán không thành công. [redacted]

optical buffering and packet-level parsing, which do make OBS become a much promised paradigm.

Nevertheless, there are still a lot of issues to deal with concerning the OBS realization, one of which is to how assemble IP packets into optical burst at the edge node in the OBS networks. A side effect incurred by such burst assembly process is to add additional delay for incoming IP packets. Therefore, the typical End-To-End (ETE) delay of a data burst mainly consists of three parts: burst assembly delay at edge nodes, offset time for reserving data burst, and the propagation delay in optical links from source edge node to destination edge node.

So far, most burst assembly algorithms can be classified as time threshold based [3], size threshold based [4] and hybrid assembly algorithms [5, 6, 7, 8]. In time threshold based assembly algorithms, a data burst is generated and issued into the OBS core networks periodically, the size of data bursts thus varies with the IP traffic load. One obvious disadvantage of time threshold based assembly is unable to control the size of data burst effectively, i.e. the size of data burst is too short at light traffic, and too long at high traffic. However, in size threshold based assembly algorithms, fixed-size burst are generated aperiodically at edge node. Similarly, there is also one of drawbacks that the time of assembling a data burst is too long at light traffic, which results in huge End-To-End delay. Thus, hybrid assembly algorithms are

[redacted]

[redacted]

proposed to overcome the aforementioned problems by using two thresholds, time and size threshold, to generate one data burst.

Although hybrid assembly algorithms overcome the shortages of time and size threshold based assembly algorithms, how to further reduce the burst assembly delay is above and beyond their capacity. In order to improve hybrid assembly algorithms, paper [9, 10] proposed a new method, Forward Resource Reservation FRR, to reduce the burst assembly delay. One attractive feature is that the burst assembly procedure and the transmission of a BHP process in parallel not in sequence, and thereby minimize their impact on total End-To-End delay. However, one shortage of FRR schemes is that bandwidth will be wasted due to unsuccessful size prediction.

In view of the drawbacks of the abovementioned schemes, we propose new scheme about burst assembly algorithm based on prediction of both size and time. Using two predictive parameters, size and time, our scheme can not only effectively control the prediction success probability, but also eliminate bandwidth wastage due to unsuccessful prediction as well. We also investigate hybrid threshold to detect the alternation of incoming IP traffic load. Based on this, the burst size is able to vary with the traffic load by step size. Furthermore, an aggressive algorithm about predictive

[redacted]

coefficient a is proposed to deliver performance improvement. This paper is organized as follow. In section 2, we introduce the basic conception of burst assembly mechanism in OBS network. In section 3, we propose burst assembly algorithm based on size and time prediction, termed as BASTP. In section 4, simulation results and analysis. In section 5, concludes this paper.

2 Concept of Burst Assembly Mechanism

Fig. 1: Basic burst assembly mechanism

(1) After using classifier to classify IP packet according to the IP traffic class and egress edge node, the IP packet is sent into corresponding buffer.

(2) When threshold, like time and size, is reached, a new burst is generated and sent into Scheduler.

(3) The scheduler use certain algorithm to set offset time, at the same time, a corresponding BCP is created and sent. The BCP carries essential information like route, data channel, class of service, offset time and so on.

(4) When the offset time is reached, the data burst (DB) is immediately sent into the OBS networks.

3 Burst Assembly Algorithm Based on Size and Assembly Time Prediction

In this section, a novel burst assembly algorithm based on burst size and assembly (BASTP) is proposed to reduce the End-To-End delay and implement adaptive assembly algorithm.

3.1 Hybrid threshold

[redacted]

[redacted]

By analyzing adaptive assembly algorithms [6, 7, 8, 11], we find that the adaptive assembly algorithms lengthen the burst size and shorten the burst assembly time as the traffic load goes up, contrariwise, shorten the burst size and lengthen the burst assembly time as the traffic load goes down. If the burst size doesn’t vary or varies a little with the traffic load, the adaptive assembly algorithms will change into size threshold based algorithm. Similarly, if the burst assembly time doesn’t vary or varies a little with the traffic load, the adaptive assembly algorithms will change into time threshold based algorithm also. Thereby, two laws can be obtained based on above analysis.

(a) When traffic load goes up, burst size is apt to lengthen with shortening burst assembly time

(b) When traffic load goes down, burst size is apt to shorten with lengthening burst assembly time

By synthesizing (a) and (b), a conclusion can be drawn:

(c) Burst size and burst assembly time make an opposite direction in change as traffic load varies.

As seen from Fig.2, we can get the trend figure of both burst size and burst assembly time as a function of traffic load. In order to further explain issue, we suppose that the IP traffic is constant not bursty and burst size, termed as S0, is fixed. To a certain traffic load, there must

[redacted]

be a corresponding burst assembly time, termed as T0, vs. S0. And the product of T0 and S0 is termed as hybrid threshold (HT). When the traffic load is going up, the burst assembly time, termed as Ti, must be shortened based on same burst size (S0) according to law (a), which make hybrid threshold (HT) is more than the product of S0 and T1. Similarly, when the traffic load is going down, the burst assembly time, termed as T2, must be made long based on same burst size (S0) according to law (b), which make HT is less than the product of S0 and T2, Therefore, the traffic load can be detected if a appropriate HT is found. Once detecting the alteration of traffic load, it is easy to implement an adaptive burst assembly algorithm. So, how to set the HT is become the key issue. If the HT is set much more or less, the conclusion (c) cannot be used to detect the alternation of traffic load. Besides, although our analysis based on constant traffic, the method to detect the alteration of traffic load is also adapted for bursty traffic because the bursty traffic is just that the traffic varies quickly and suddenly. Consequently, using hybrid threshold to detect the alternation is totally different from the schemes have been proposed.

Traffic. Load

Fig. 2: burst size and burst assembly time vs. traffic load

Fig. 3: Step size vs. number of step

3.2 BASTP Scheme

Based on the abovementioned analysis, a new burst assembly algorithm is proposed with the feature that both burst size and burst assembly time is predicted. The size from the minimum burst size to maximum is divided into N equal part that is termed as step size. And, the burst size is adjusted by the step size according to the alternation of step. We can also see that from Fig.3. A hybrid threshold, the product of burst size and burst assembly time, is preset to detect the alternation of traffic load. While first IP packet flow into assembly buffer, the BCP, carrying the predictive burst size and offset time etc., is created and sent for reserving link resources for corresponding unfinished data burst.

The time order of prediction can be seen from Fig.4. A BASTP scheme involves a three-step phases as follows:

Phase One: Pretransmission. When the first IP packet arrives at the buffer, the BCP is constructed instantly upon the prediction, calculated by the previous data burst. Other necessary information about route, wavelength etc. is also injected into the BCP. Once completing the creation of BCP, the BCP will be launched into the OBS network at Th + t, where t is used in selecting route and assigning wavelength and so on.

Fig. 4: Time order figure of predictive assembly algorithm

Phase Two Completion Of Assembly. When the predictive burst size or burst

[redacted]

assembly time is reached, the data burst is completed assembly.

[redacted]

Phase Three Prediction. Once the data burst is completed, the actual burst size and assembly time is used to predict both burst size and assembly time of the incoming data burst.

Fig. 5: Flow chart of BASTP scheme

The flow chart of BASTP assembly algorithm can be seen from Fig.5. The main step of BASTP is showed as follow:

(1) Initialization

Lequal is equal part; Lmax is the maximum burst size; Lmin is the minimum burst size; N is the numbers of step.

[redacted]

(2)

Lstep is the current step size (Lmin < Lstep < Lmax); step is the current number of step.

where To is the offset time; Tfa is the predictive assembly time; Ts is the time used in scheduler.

Counter, used to record the size of the assembling burst, is set 0; Timer, used to record the elapsed time of the assembling burst, is set 0.

(2) If the IP packet, received by buffer, is the first packets, the Counter and Timer is started-up, meanwhile, the pretransmitted BCP is created and issued into OBS networks for reserving link resource. Both predictive burst size and the offset time, carried in the BCP, are gotten according to equation (2) and (3) respectively.

[redacted]

(3) When the counter reaches the

[redacted]

predictive burst size (Lstep) or the timer reaches the predictive burst assembly time (Tfa), the data burst complete the assembly, and is sent into scheduler. At the same time, the actual assembly time (Ta) is used to update the predictive assembly time (Tfa) of the next incoming data burst according equation (4), then go to step (4) (4) where Tfa is predictive assembly time; Ta is the actual assembly time; T(i) is the is assemble time of the i-th data burst, and the assembly time of the next incoming burst is then predicted according to those of the previous M bursts.	[redacted]

(4) Compare the hybrid threshold with the product of Ta and Lstep.
if Ta x Lstep < HT then
if TCounter + + > Upper-Limit then TCounter = 0; step = step + 1
else
if TCounter < Low-Limit then
TCounter = 0; step = step — 1
Where TCounter is used to record the time that the product of Ta and Lstep is more than or less than HT; HT, the hybrid threshold, is used to detect the alternation of traffic load.
(5) Once completing comparison, both Counter and Timer are reset 0 in order to make the preparation for receiving the next incoming IP packet.
	[redacted]

3.3 Predictive coefficient a Whether a data burst is completed assembly is depend on two thresholds, Lstep and Tfa. However, different threshold is reached has different effect on the predictive success probability,	[redacted]

Termed as PSP. If Lstep is reached, it represents that the predictive burst size is equal to actual burst size; namely, the predictive success probability (PSP) is 100 percent. If Tfa is reached, it represents that the predictive burst size is more than the actual burst size, namely, the PSP falls short of 100 percent.

By enlarging Tfa artificially, Lstep is going to have more opportunity to be reached, namely, the PSP can be controlled by introducing small redundant time. Thus, a predictive coefficient,
termed as a, is proposed to offer such small redundant time, which makes the equation (4) become equation (5).
(5)

At the same time, predictive coefficient a is also directly relative to the End-To-End delay, so, by adjusting a appropriately, the performance, like End-To-End delay and PSP, can be improved according to the alternation of traffic load. For example, if traffic load is light, a is going to be diminished so that the too long End-To-End delay can be controlled. Although small a will cause PSP fall, the End-To-End delay vs. PSP is apparently important at light traffic load. Contrariwise, if traffic load is heavy, a is going to be raised so that the predictive success probability can be up. High predictive success probability means high bandwidth utilization because predictive success probability is proportion of actual burst size to the predictive burst size, namely, the burst size carried in the pre-transmitted BCP.

[redacted]

[redacted]

In order to deliver better performance improvements, a is planned to adjust according to the adjustment of Lstep. Since Lstep varies with the alternation of traffic load, predictive coefficient a is going to be adaptive also. The value of a is going to be diminished while the traffic load goes down, vice versa.

3.4 Features

From the above description, the burst assembly algorithm based on burst size and assembly time prediction (BASTP) has the following key features:

- Low Latency: the End-To-End delay can be diminished to the least extent because the predictive BCP is sent into OBS network for the corresponding unfinished data burst while the first IP packet is coming into assembly buffer.
- Both Predictions: besides size prediction, the assembly time prediction is introduced. If any prediction condition is reached, a burst will be finished assembly. Thus, one advantage of both predictions is to eliminate the situation that the predictive BCP is deemed a failure if the actual burst size exceeds the predictive burst size.
- Adaptability: since the alternation of traffic load can be detected by using hybrid threshold, it is easy to implement that the burst size and assembly time vary adaptively.

- Predictive Success Probability Controllable: predictive coefficient a is proposed to adjust the End-To-End delay and predictive success probability. In order to better deliver performance improvement, the adaptive a is used to

[redacted]

shorten the End-To-End delay at light traffic load, and raise the predictive success probability at heavy traffic load.

4 Simulation and Analysis

In this section, we compare the overall performance of the proposed BASTP algorithm versus other typical burst assembly algorithms. A time-driven and event-driven Optical Burst Switching Networks Simulation Mainframe (OBSNSM) was developed using JAVA language. Three burst assembly algorithm programs were developed, one for the BASTP algorithm and the others for time-threshold, and size-threshold algorithm with identical network parametersthe performance parameters used here are end-to-end (ETE) delay, burst size and burst drop probability.

checked

To compare the performance results of those assembly algorithms, we used a 6-nodes topology structure (Fig.6) for simulation, where the numbers on the links in the figure represent the length of links; unit is kilometer (km). In this OBS network, there are three edge nodes: node 1, 2 and 3, and three core nodes: node 4, 5 and 6. The network load is calculated by the following formula.

Network-Load — Actual Bandwidth/ EE (Bandwidth X WT) (6)

where ActualBandwidth represents the bandwidth is used to generate IP traffic, ET represents the total of edge node; and LT is total of export link of the edge node; Bandwidth means the bandwidth of one wavelength; W T is the total of wavelength of one link.

[redacted]

[redacted]

The following assumptions are also made in simulation: • IP packets arrive at an edge node following ON/OFF model with a Pareto distribution; • Packets length are exponentially distributed and the average length is 1024 Bytes; • Each core node has no wavelength conversion capability and Fiber Delay Line (FDL) • Each link has bi-directional fibers and each fiber has 8 data wavelengths and 1 control wavelength • Bandwidth of wavelength is 10 Gb/s. Wavelength assignment policy is Rand-Fit; • TS, the time of setup BCP, is set as 5 microseconds, Toxc, time of configuring OXC, is set as 15 microseconds; • Lmin is set as 500 Kbytes and Lmax is 2000 Kbytes. 4.1 Numerical results of hybrid threshold (HT) How to set a hybrid threshold is a key issue for BASTP algorithm. If HT is set inappropriately, it will not only result in the alternation of traffic load can not be detected, but also fail to implement Fig. 7: End-To-End delay vs. network load for various HT the adaptability of both burst size and assembly time. Thus, the foremost step is to find the right HT. The results can be seen from Fig.7 and Fig.8. Fig. 8: burst size vs. network load for	[redacted] [redacted] [redacted]

various HT

Three scenarios, HT=4.096E8 and HT=8.192E8 as well as HT=1.229E9, are used to compare the performance. Fig.7 compares the average End-To-End (ETE) delay as a function of network load for the three scenarios. Fig.8 presents the same comparison as Fig.7 for average burst size.

Fig. 9: End-To-End delay vs. network load for various a

As seen from Fig.7, the larger the HT, the lesser the End-To-End delay, i.e. the End-To-End delay is in inverse proportion to HT. Thus, the larger HT is inclined to be selected as long as the HT can implement the adaptability of burst size. In Fig.8, we can find the least HT cause the burst size ascends the maximal burst size too early, on the contrary, the largest HT cause the burst size fail to ascend the maximal burst size although the network load is heavy. Therefore, how to set the HT is tradeoff based on Fig.7 and Fig.8. In BASTP scheme, the HT is set as 8.192E8.

4.2 Numerical results of predictive coefficient a

Predictive coefficient a is aggressive way to adjust both the End-To-End delay and predictive success personality for performance improvement. Four scenarios, a = 0.0, a = 0.025, a = 0. 05 and Adaptive, are used to compare performance of predictive coefficient based on identical simulation parameters, where Adaptive scenario means that a varies from 0.0 to 0.05 according to the alternation of Lstep, The results can be

[redacted]

[redacted]

seen from Fig.9 and Fig.10.

Fig. 10: predictive success probability vs. network load for various a

Fig.9 compares the average End-To-End delay as a function of network load for the four scenarios. We can see from the Fig.9 that the End-To-End delay is in direct proportion to a, and, the scenario of adaptive a can vary well between the scenarios of a = 0.0 and a = 0.05.

Fig.10 presents the same compares as Fig.9 for average predictive burst size. In Fig.10, we can find that the predictive success probability is also in direct proportion to a. Since both End-To- End delay and predictive success probability are in direct proportion to the, the selection of a is inclined to be a compromise. Besides, in order to improve the End-To-End delay at the light traffic load and raise the predictive success probability for bandwidth utilization at the heavy traffic load, adaptive a is preferred in BASTP algorithm.

4.3 Numerical results of BASTP

After finding the appropriate HT and adaptive a, two scenarios, size threshold and time threshold, are selected in order to compare the performance of BASTP. The parameters of three scenarios are designed as follow

- Time threshold is set as 1000 microseconds.

Size threshold is set as 1500 Kbytes.

X. Jiang et al. /Journal of Computational Information Systems 9: 2 (2013) 463-475

- HT is set as 8.912E8; adaptive a scheme is selected.

[redacted]

[redacted]

Fig. 11: ETE delay vs. network load for various assembly algorithms

Fig. 12: burst size vs. network load for various assembly algorithms

Fig.11 shows the average End-To-End delay as a function of network load for three scenarios. It can be easily observed that the ETE delay of BASTP scenario is far less than time threshold scenario except that the network load is 0.1, and the ETE delay is also less than size threshold scenario when the network load is less than 0.6. However, the burst size of BASTP assembly algorithm is also much more than the size threshold scenario while the load is over 0.6, this can be seen from Fig.12. Thus, the ETE delay of BASTP is actually less than the size threshold scenario during the whole simulation. The lower ETE delay can be obtained not only because the BASTP is adaptive assembly algorithm but also because the predictive mode.

Fig.12 shows the average burst size vs. network load for the three scenarios. It evidently demonstrates the size threshold scenario is fixed while the time threshold scenario increases linearly as the network load goes up. When the network load is over 0.4, time threshold scenario creates bigger DB than the other two. Obviously, BASTP scenario is tradeoff relative to size and time threshold scenarios. Besides, since the burst size of BASTP scenario is adjusted by step size (Lstep), the BASTP is apt for the slot scheme [12, 13] with little adjustment.

[redacted]

Fig.13 shows the average burst drop probability as a function of the network load. As the network load increase, the burst drop probability of BASTP scenario increases, and the burst drop probability of BASTP scenario is similar to the others. Since no special methods or devices, like reroute, Fiber Delay Line and so on, are used to resolve the collisions, the burst drop probability is not improved so obviously by the BASTP scenario.

Fig. 13: burst drop probability vs. network load for various assembly algorithms

5 Conclusions

In this paper, a novel BASTP assembly algorithm has been proposed and proved to be practical in reducing the ETE delay and implementing the adaptability of assembly algorithm. The BASTP assembly algorithm consists of four features: a predictive BCP, both size and assembly time prediction and controllable predictive success probability as well as hybrid threshold. Furthermore, traditional time threshold and size threshold assembly algorithms are investigated, and a novel burst assembly algorithms, namely BASTP assembly algorithm is proposed for OBS. Using simulation, we find the appropriate hybrid threshold (HT) to detect the alternation of traffic load, and get the adaptive predictive coefficient algorithm. Based on these simulations, the results of three assembly algorithms for ETE delay, burst size and burst drop probability, are obtained. The simulation results indicate that the

[redacted]

proposed assembly algorithm will not only take current traffic load into account to adjust burst assembly adaptively, but also reduce the latency obviously. With the proposed BASTP assembly algorithm, two issues are found to deserve further study in the future: • How to adjust hybrid threshold (HT) adaptively. Since the HT is constant from the light network load to heavy load, the HT is not optimized correspondingly according to the different network load, and the HT is set manually not adaptively. Thus, an algorithm, to adjust HT adaptively, may be an effective way to improve the performance further.	[redacted]