www.ietdl.org

Published in IET Communications
Received on 9th September 2009
Revised on 17th April 2010
doi: 10.1049/iet-com.2009.0587

ISSN 1751-8628

Adaptive dynamic bandwidth allocation algorithm with sorting report messages for Ethernet passive optical network

W.-P. Chen[1] W.-F. Wang[2] W.-S. Hwang[1]
[1]Department of Electrical Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung 807, Taiwan
[2]Department of Information Engineering, National Yunlin University of Science and Technology, Yunlin 64002, Taiwan
E-mail: wwf@yuntech.edu.tw

Abstract: Broadband access networks using passive optical network (PON) technology can extend the transmission distance and increase the transmission capacity of carrier networks. One PON solution, the Ethernet passive optical network (EPON), can provide huge bandwidth capacity, low cost, simple architecture and easy maintenance. Therefore EPON becomes a promising candidate for future last-mile solutions. To prevent data collision and ensure efficient transmission, EPON must employ a media access control (MAC) protocol to allocate the shared resource of a common upstream transmission medium. This article proposes a novel DBA algorithm that sorts all REPORT messages by the request length at the next transmission cycle to fully utilise the idle time between cycles as long as at least one optical network unit (ONU) requests a long enough transmission window. Alternately, when no grant length is long enough, then some of ONUs' requests are laid out together in the idle period to utilise the otherwise wasted idle time. Event-driven simulations show that Sort-DBA can significantly improve the network performance in terms of packet delay, average queue length and throughput, as compared with the well known IPACT, DBA2 and scheduling control scheme DBA algorithms.

1 Introduction

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. This tremendous growth of Internet traffic has aggravated the lack of access networks. The 'last-mile' between local area networks (LANs) and metropolitan area networks (MANs) remains the main bottleneck. The most widely deployed broadband solutions today are digital subscriber line (DSL), community antenna television (CATV) and cable modem (CM)-based networks. These networks, however, do not provide enough bandwidth to support the growing demand. In addition, these networks have a limitation that the distance of any xDSL subscriber to a central office (CO) must be less than 1800 feet because of signal distortion. Therefore a next-generation broadband access network must provide not only increased bandwidth at low cost to end users, but also must provide service for long-distance access networks.

Passive optical networks (PONs) [1–14] have aroused interest from both industry and academia as a feasible and cost-effective solution. A great deal of effort has gone into developing and standardising various PON technologies. EPONs [7–14] represent the convergence of low-cost Ethernet equipment (e.g. switches and network interface cards) and optical fibre architecture. Considering that more than 90% of data traffic originates from and terminates in Ethernet LANs in this moment, EPONs appear to be a natural choice for future last-mile solutions. This technology is currently being developed and standardised by the IEEE 802.3ah task force [14] in the hope of significantly improving broadband service while minimising equipment, operation and maintenance costs.

doi: 10.1049/iet-com.2009.0587

An EPON is a point-to-multipoint (P2MP) network consisting of one optical line terminal (OLT) and multiple optical network units (ONUs). The OLT broadcasts to all ONUs simultaneously in the downstream direction, whereas in the upstream direction, a single optical fibre channel is shared by all ONUs. To avoid data collision, a scheduling algorithm is needed to prevent simultaneous transmissions. The well-known media access control (MAC) protocol CSMA/CD is a standard for Ethernet LANs [8], but does not represent a good choice for EPONs. Since the OLT will receive all data packets transmitted by the ONUs and discard those involved in collisions, each ONU would require an additional receiver operating at the upstream wavelength and a carrier sensing circuit. This solution would greatly increase the network cost. Besides, its bandwidth utilisation is extremely low because the collision packets make too many data retransmissions especially under heavy traffic load conditions.

Time division multiplexing (TDM) technology is a popular alternative for EPONs. Each ONU is assigned a timeslot for data transmission in each cycle and can only transmit data in the allocated window. It can be either static or dynamic, depending on the arbitration mechanism implemented by the OLT. Kramer and Pesavento [9] studied the performance of EPON under TDM using a fixed bandwidth assignment algorithm. Although this scheme is easy to implement and performs well under heavy load conditions, it cannot handle statistical multiplexing between ONUs. Static schemes based on TDM are also very inefficient because of the bursty nature of access network traffic.

To cope with this problem, Kramer *et al.* [10] proposed a polling-based scheme called interleaved polling with adaptive cycle time (IPACT). This algorithm achieves good performance by combining limited service with a maximum transmission window defined over 2 ms polling cycles. However, the idle time issue is not effectively resolved and the fact that IPACT allocates bandwidth based on a single ONU REPORT is not globally optimised. Owing to the bursty nature of Ethernet traffic and encapsulation of Ethernet packets (i.e. packet fragmentation is prohibited according to IEEE 802.3 [14]), some ONUs may have less traffic to transmit, whereas other ONUs may have more traffic to transmit and need more bandwidth in each transmission cycle.

To address the issue, Luo *et al.* [11] proposed a DBA scheme called limited sharing with traffic prediction (LSTP) that predicted the arriving traffic during the waiting time and maintained a portion of the bandwidth for delivery. However, the prediction scheme has the behaviour of a bursty traffic, so some bandwidth may be wasted because the scheme cannot estimate accurately the real traffic load demand for all ONUs at the next transmission cycle. Assi *et al.* [12] proposed a DBA algorithm, which utilises the excessive bandwidth of lightly loaded ONUs to carry some of the bandwidth demand of heavily loaded ONUs in each transmission cycle, thus improving the performance of the limited allocation scheme. In addition, also addressing the idle time issue, the authors proposed an early allocation mechanism, called DBA2, which schedules a lightly loaded ONU without delay, whereas it schedules heavily loaded ONUs after the OLT receives all REPORT messages and performs computation for bandwidth allocation. However, the DBA2 algorithm improves the idle period only under light or medium traffic loads. Moreover, most of the ONUs may have a bandwidth demand larger than the minimum guaranteed bandwidth under high traffic loads, so the GATE message cannot be transmitted early to the ONU for idle time compensation. In 2006, Zheng [13] proposed a bandwidth allocation called new scheduling control that uses a tracker value to address the idle time problem under high traffic loads. Although this algorithm improves the DBA2 idle time issue under heavy load conditions, it still wastes bandwidth under heavy load conditions because of the redundant overheads of the processing time of the tracker and the regular REPORT messages. Besides, it has an unfairness issue such that when the previous ONUs are operating under a light load, the following ONUs can share the remainder bandwidth, but the previous ONUs cannot allot the remainder bandwidth if the following ONUs are operating under a light load.

From the above, to improve bandwidth utilisation so as to address the idle time issue under medium or heavy traffic load conditions, we propose a new DBA algorithm called the Sort-DBA algorithm in which the transmission order of all grant data are allowed by the REPORT length at the next cycle time. The goal of this transmission scheme is to minimise the idle time under any traffic load conditions. In fact, the Sort-DBA algorithm can completely eliminate the idle time between cycles as long as at least one ONU has a sufficiently long data transmission. Even when most of the ONUs are operating under a medium load, with the waste of some guard-band distance, the new algorithm can still achieve good results with regard to eliminating idle time. Moreover, this paper also proposes a queue management method, which reduces the unnecessary overhead of REPORT messages when the ONUs are operating under heavy traffic, thereby achieving higher bandwidth utilisation under heavy traffic in EPON systems. It will be shown in our simulation experiment that Sort-DBA has better bandwidth utilisation, lower delay and lower queue length than IPACT [9], DBA2 [11] and the scheduling control scheme DBA [12] on EPON systems.

The remainder of the article is organised as follows. Section 2 summarises the basic EPON architecture. The new Sort-DBA algorithm supporting EPON is detailed in Section 3. Section 4 discusses the assumptions behind our system simulation and describes the results of simulation comparing Sort-DBA to IPACT, DBA2 and scheduling control DBA algorithms. Section 5 concludes with a few remarks.

2 EPON architecture

There are several multipoint topologies suitable for an EPON: bus, ring, tree and tree-and-branch. The most popular choice is based on a star topology. As shown in Fig. 1, it consists of one optical line terminal, a 1:N passive star splitter (and combiner) and multiple ONUs. The number of ONUs (N) is typically between 4 and 64, but networks with $N = 128$ have also been fabricated. The OLT resides in a CO that connects the access network to a metro core network or wide area network (WAN). The OLT is connected to a passive star splitter by a single optical fibre. The passive splitter is generally located far from the CO, but close to the subscriber premises. An ONU may be located at a curb or building, or even on the subscriber premises, and is connected to the passive splitter by a short, dedicated optical fibre. The distance between the OLT and an ONU typically ranges between 10 and 20 km. Presumed to be compatible with the IEEE 802.3 standard, all data is encapsulated in Ethernet packets for transmission. The fragmentation of Ethernet packets in the transmission window is not allowed. All transmissions occur between the OLT and the ONUs.

In the downstream direction, the OLT connects all ONUs as a point-to-multipoint (P2MP) architecture. It broadcasts Ethernet frames to all ONUs simultaneously through the 1:N splitter on a single wavelength (e.g. 1550 nm). This behaviour is similar to that of a shared media network. In the downstream direction, the Ethernet standards fit the EPON architecture perfectly: packets broadcast by the OLT are given a MAC address, so they will be extracted only at the intended destination (that is, an ONU). In the upstream direction, an EPON is a multipoint-to-point (MP2P) network. All ONUs transmit their data to the OLT on a common wavelength (e.g. 1310 nm) through the 1:N passive combiner. Since the ONUs share the upstream transmission medium, an EPON must efficiently allocate uplink access and avoid data collisions. A MAC-based mechanism is generally chosen for this purpose.

The multipoint control protocol (MPCP) [15] has been widely used to implement DBA in EPONs. MPCP is a signalling protocol currently being developed and standardised by the IEEE 802.3ah task force. At the moment, MPCP does not specify a particular bandwidth allocation algorithm. Rather, it provides an effective control mechanism, which facilitates the implementation of bandwidth allocation algorithms. MPCP has two operation modes: normal and auto-discovery. In the normal mode, MPCP relies on GATE and REPORT Ethernet control messages to allocate bandwidth. A GATE message is used by the OLT to allocate a transmission window. REPORT messages are used by ONUs to communicate their local conditions to the OLT. In its auto-discovery mode, the protocol relies on three control messages: REGISTER, REGISTER_REQUEST and REGISTER_ACK. These are used to discover and register a newly connected ONU and to collect related information such as round-trip time (RTT) and its MAC address.

3 Sort-DBA algorithm for EPON systems

In a global polling-based DBA algorithm, each ONU sends normally a REPORT message to the OLT after transmitting its grant data. The purpose of these messages is to request bandwidth for the next cycle time. Once the OLT has received all REPORT messages, it calculates the appropriate bandwidth allocation and broadcasts GATE messages to all ONUs while the upstream bandwidth is idle between the time of the last REPORT message (R_N) is received and the time the GATE message (G_1) is sent to

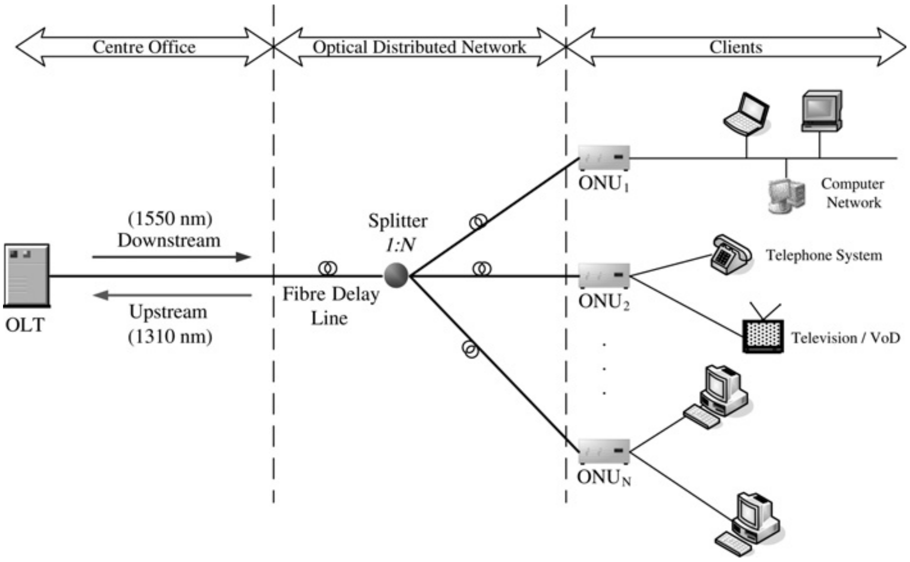

Figure 1 *EPON star architecture*

doi: 10.1049/iet-com.2009.0587

ONU$_1$. From Fig. 2, the idle time is given by

$$T_{\text{idle}} = T_{\text{dba}} + \text{RTT} + T_{\text{ONU}} \tag{1}$$

where RTT is the round-trip time from ONU to OLT, T_{dba} is the processing time of the DBA algorithm and T_{ONU} is the processing time of the ONU (on receiving a GATE message).

In terms of bandwidth utilisation, the idle time defined above is undoubtedly wasted. To address this problem, we propose a novel DBA algorithm by sorting all REPORT messages with the request length to give a transmission order. This DBA algorithm can be divided into two cases to eliminate the idle time according to the bandwidth demands of all ONUs. The two cases are described as follows

3.1 At least one REPORT is long enough

In fact, the idle time issue of a global polling-based DBA algorithm occurs because the OLT must wait for the last REPORT message before executing the DBA algorithm and then transmitting GATE messages to ONU$_1$. Thus, if the last REPORT message (R_N) can be transmitted before the last grant data (L_N) and if the L_N is long enough to compensate for the idle period of formula (1), then the idle period can be completely eliminated. To ensure low delay, other grant data must be transmitted before the REPORT messages and the transmission order must be allowed by the grant length. Hence, to minimise the idle time, the minimum guaranteed bandwidth of the last granted bandwidth per cycle ($L_{\min}$) should be at least

$$L_{\min} = R_{\text{u}}(T_{\text{idle}} - T_{\text{g}}) \tag{2}$$

where R_{u} is the total upstream bandwidth of the fibre and T_{g} is the guard-band time between two neighbouring packets.

For example, in Fig. 3, the last REPORT message (R_{N-1}) received by the OLT before the last grant data (L_{N-1}) at cycle time I can use up some of the idle time that would normally occur before cycle time ($I+1$). During a DBA calculation time, the OLT sorts all of the request lengths in order to find the size of request lengths R_1, R_N, $R_4, \ldots$ and R_3; furthermore, R_1 is long enough to compensate for the idle period at cycle time ($I+1$). The OLT transmits a series of GATE messages (G_N, $G_4, \ldots,$ G_3, G_1) to all of the ONUs. Since the length of the granted bandwidth (L_{N-1}) is long enough at cycle time I, the OLT is sufficient time to execute the DBA algorithm and transmit GATE messages to all of ONUs under the idle period. In this way, the idle time problem can be neatly solved as long as at least one ONU has a sufficiently long data transmission.

If all of the ONUs are under heavy traffic load, the maximum cycle time ($T_{\max}$) is given by formula (3), where

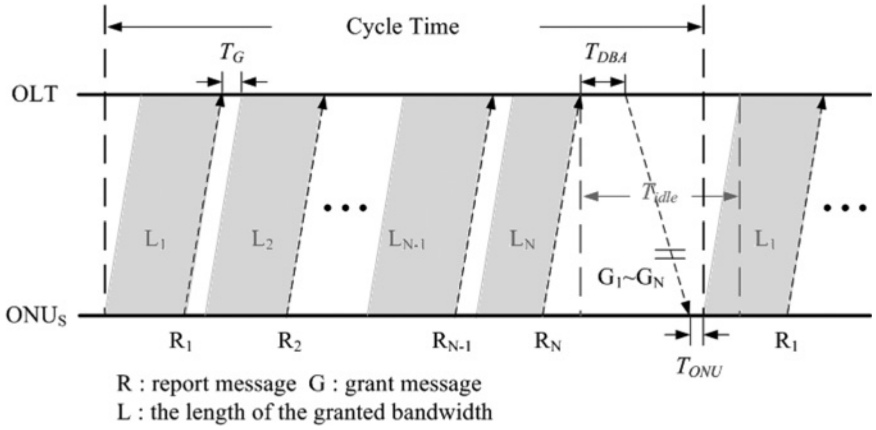

Figure 2 *Idle time issue in an EPON polling scheme*

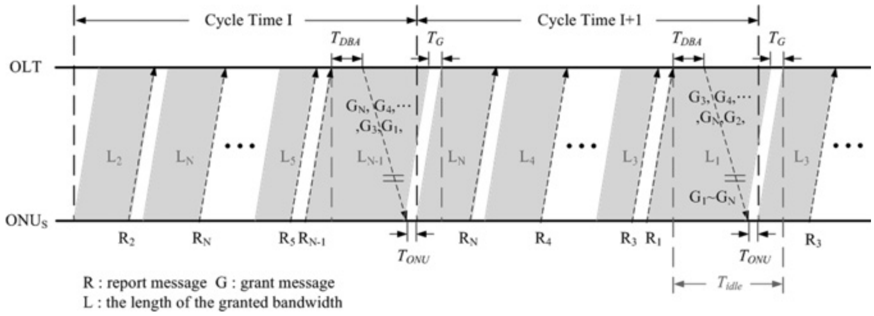

Figure 3 *Example for the Sort-DBA algorithm under high traffic load conditions*

doi: 10.1049/iet-com.2009.0587

H is the length of an ONU REPORT frame and K_i $(=1, 2, 3, \dots)$ is the multiple of $L_{\min}$ assigned to each ONU, based on the contract of service level agreement (SLA)

$$T_{\max} = \sum_{i=1}^{N} \left(\frac{H + K_i L_{\min}}{R_u} + T_g \right) \tag{3}$$

Therefore the maximum bandwidth utilisation ($\eta_{\max}$) and the maximum throughput ($P_{i,\max}$) of each ONU$_i$ can be obtained as (see (4))

$$\begin{aligned} \eta_{\max} &= \frac{K_1 L_{\min} + K_2 L_{\min} + \cdots + K_N L_{\min}}{(H + T_g R_u + K_1 L_{\min}) + (H + T_g R_u + K_2 L_{\min}) + \cdots + (H + T_g R_u + K_N L_{\min})} \\ &= \frac{\sum_{i=1}^{N} K_i L_{\min}}{N(H + T_g R_u) + \sum_{i=1}^{N} K_i L_{\min}} \end{aligned} \tag{4}$$

$$\begin{aligned} P_{i,\max} &= R_u \times \frac{K_i L_{\min}}{N(H + T_g R_u) + (K_1 + K_2 + \cdots + K_N) L_{\min}} \\ &= \frac{R_u}{N} \times \frac{K_i L_{\min}}{(H + T_g R_u) + (L_{\min}(K_1 + K_2 + \cdots + K_N)/N)} \\ &= \frac{R_u}{N} \times \frac{K_i L_{\min}}{(H + T_g R_u) + (1/N) \sum_{i=1}^{N} K_i L_{\min}} \end{aligned} \tag{5}$$

Under uniform traffic load conditions, that is $K_1 = K_2 = \cdots = K_N = 1$, formulas (3)–(5) can be expressed as

$$T_{\max} = N \left(\frac{H + L_{\min}}{R_u} + T_g \right) \tag{6}$$

$$\eta_{\max} = \frac{N L_{\min}}{N(H + L_{\min} + T_g R_u)} = \frac{L_{\min}}{H + L_{\min} + T_g R_u} \tag{7}$$

$$P_{i,\max} = \frac{R_u L_{\min}}{N(H + L_{\min} + T_g R_u)} \tag{8}$$

Formula (7) shows that the maximum bandwidth utilisation is unrelated to the number of ONU, that is bandwidth utilisation scales are perfect in this case.

Owing to the bursty nature of access network traffic, however, the ONU will waste overhead with their REPORT messages under heavy traffic load conditions. Fortunately, an effective queue management scheme can overcome this problem. First, we assume that each ONU has a queue whose maximum capacity for incoming packets is $m \times L_{\min}$. If the current arrival queue length of ONU$_i$ is $x \times L_{\min}$ $(m \geq x)$, then ONU$_i$ notifies the OLT in the REPORT frame. The OLT then replies with the GATE frame to inform ONU$_i$ that it need not transmit further REPORT messages for the next $(x-1)$ cycles. By this means, considerable savings can be achieved when all the ONU are operating under heavy traffic load conditions. We can rewrite the maximum bandwidth utilisation and throughput of each ONU at uniform traffic load condition as

$$\begin{aligned} \eta_{\max} &= \frac{x N L_{\min}}{N(H + L_{\min} + T_g R_u) + N(x-1)(L_{\min} + T_g R_u)} \\ &= \frac{x L_{\min}}{H + x(L_{\min} + T_g R_u)} \end{aligned} \tag{9}$$

$$P_{i,\max} = \frac{R_u}{N} \times \frac{x L_{\min}}{H + x(L_{\min} + T_g R_u)} \tag{10}$$

By comparing formula (7) with formula (9), we can obtain an expression for the enhancement to maximum bandwidth utilisation because of queue management on a uniform traffic system

$$\begin{aligned} D &= \frac{x L_{\min}}{H + x(L_{\min} + T_g R_u)} - \frac{L_{\min}}{H + L_{\min} + T_g R_u} \\ &= \frac{1}{1 + (T_g R_u + H x^{-1}/L_{\min})} - \frac{1}{1 + (T_g R_u + H/L_{\min})} \end{aligned} \tag{11}$$

3.2 No granted bandwidths are long enough

In the above method, the idle time can be completely solved if at least one REPORT is long enough ($L_{\min}$) in the next cycle time. A corollary is that if all the ONUs are operating at light traffic load, the OLT will be unable to find any one REPORT to fill the idle time. To address this issue, a compensation scheme is proposed so that other grant data will be laid out together during an idle period after the last granted bandwidth so as to use the remainder space of the idle time (designated L_{ins}). A guard-time distance (L_g) is necessary between each neighbouring granted bandwidth of the ONUs, which will be laid out together in the L_{ins}. Therefore to minimise bandwidth waste when using the remainder L_{ins}, each select granted bandwidth must be larger than L_g.

The processing steps of the compensation scheme are described in the following. It is assumed that the maximum length of the granted bandwidth is smaller than $L_{\min}$ and the sizes of all granted bandwidths have already been sorted by the OLT as ONU$_{N-1}$, ONU$_2$, ONU$_N$, … , ONU$_1$, ONU$_5$ at cycle time $(I+1)$, that is $L_{N-1} \geq L_2 \geq L_N, \dots, \geq L_1 \geq L_5$.

1. The maximum granted bandwidth of ONU$_{N-1}$ is still designed to be the last ONU in cycle time $(I+1)$. At the moment, OLT calculates the size of L_{ins} $(= L_{\min} - L_{N-1})$. If

doi: 10.1049/iet-com.2009.0587

the L_{ins} is larger than L_{g}, OLT starts to search for a suitable length of granted bandwidth from the other ONU_i into which it can insert the L_{ins}. The DBA algorithm, on the other hand, ended the compensation scheme if the L_{ins} is small. Here, to decrease the complexity of the compensation scheme and avoid increasing transfer delay, the search of ONU_i is from the minimum length of the granted bandwidth (L_5).

2. If L_5 is larger than L_{g}, then L_5 is selected and the OLT recalculates the length of the L_{ins} ($= L_{\mathrm{ins}} - L_5 - L_{\mathrm{g}}$).

3. If the remainder length of L_{ins} is still larger than L_{g}, the DBA algorithm repeats step (2) to find the next grant data for compensating the L_{ins}. However, when the second maximum request length of ONU_i (ONU_2) has been searched or the remainder L_{ins} is smaller than L_{g}, then OLT terminates the DBA algorithm.

In the scheme, the maximum number of searches for suitable request data is at most $N-1$ times. Note that the compensation scheme may not be able to find any suitable request data to stuff in the L_{ins} from the minimum request data after the second maximum request data has been performed. The reason is that all of ONUs are operating under light traffic load, so the L_{ins} cannot be compensated. However, under light traffic conditions, the transfer delay of all request data is very short. Under the uniform and light traffic load conditions, the minimum cycle time (T_{min}) of the Sort-DBA algorithm can be obtained as

$$\begin{aligned} T_{\mathrm{min}} &= \sum_{i=1}^{N}\left(\frac{H_i}{R_{\mathrm{u}}} + T_{\mathrm{g}}\right) + T_{\mathrm{dba}} + \mathrm{RTT} + T_{\mathrm{ONU}} \\ &= N\left(\frac{H_i}{R_{\mathrm{u}}} + T_{\mathrm{g}}\right) + T_{\mathrm{idle}} \qquad (12) \end{aligned}$$

Fig. 4 illustrates an example of the compensation scheme. At cycle time I, the OLT is starting to sort all of the REPORT messages by length, after receiving the last REPORT (R_2) at T_0. If the maximal REPORT message (G_Y) is not long enough to eliminate the idle time at cycle time $(I+1)$, the OLT begins to search the suitable grant data from the minimal request length (G_X) to insert into L_{ins}. During the processing time, assume that the request lengths of G_K and G_N are searched for suitable candidate grant data to insert into the L_{ins} by the above processing steps of the compensation scheme. A set of sequence numbers (G_{N-1}, $G_1, \ldots, G_X$, G_K, G_N, G_Y) messages will be transmitted to notify the transmission time and grant data at T_1 in the cycle $(I+1)$ for all ONUs. And then the (G_K, G_N) messages will also be transmitted to notify the ONU_K and ONU_N at T_2. In this scheme, both ONU_K and ONU_N have been received the GATE messages two times; the former only records the transmission time of the REPORT message and the latter sends a grant message including data size and transmission time in the cycle time $(I+1)$.

From the above, the flowchart of the Sort-DBA algorithm can be presented as in Fig. 5.

4 Simulation results and discussion

In this section, we evaluate the performance of our proposed DBA algorithm by a comparison with the existing DBA algorithm proposed in [8, 10, 16] based on simulation experiments. For ease of discussion, the IPACT with fixed service, the IPACT with limited service and scheduling control scheme DBA algorithms are abbreviated as IPACT (fixed), IPACT (limited) and SC-DBA, respectively. In the evaluation, we use average transfer delay, queue length, throughput and grant length per ONU as performance metrics. The transfer delay is defined as the time from a packet arriving to an ONU queue until it is received by the OLT. The throughput is defined as the total amount of traffic (in bits) delivered per second normalised by the line rate of the upstream channel. The grant length represents the average grant data length per ONU during a transmission cycle time.

4.1 Simulation assumption and model

The performance metrics are assumed to be as follows:

(a) The total number of ONUs is N. They share an upstream bit-rate of R_{u} bps.

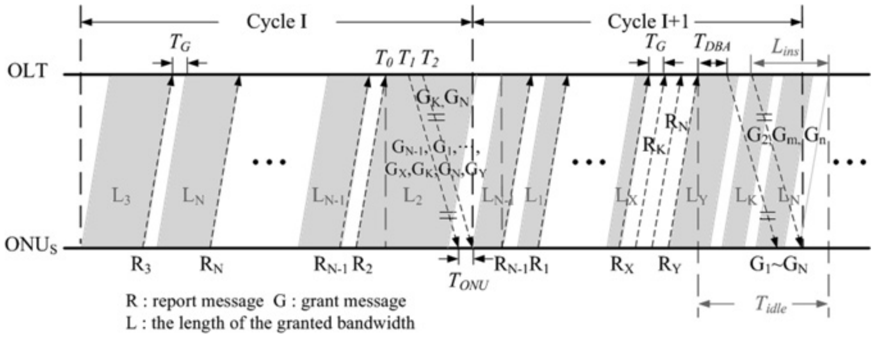

Figure 4 *Example of the Sort-DBA algorithm's compensation scheme when the EPON system is operating at medium load in transmission cycle (I + 1)*

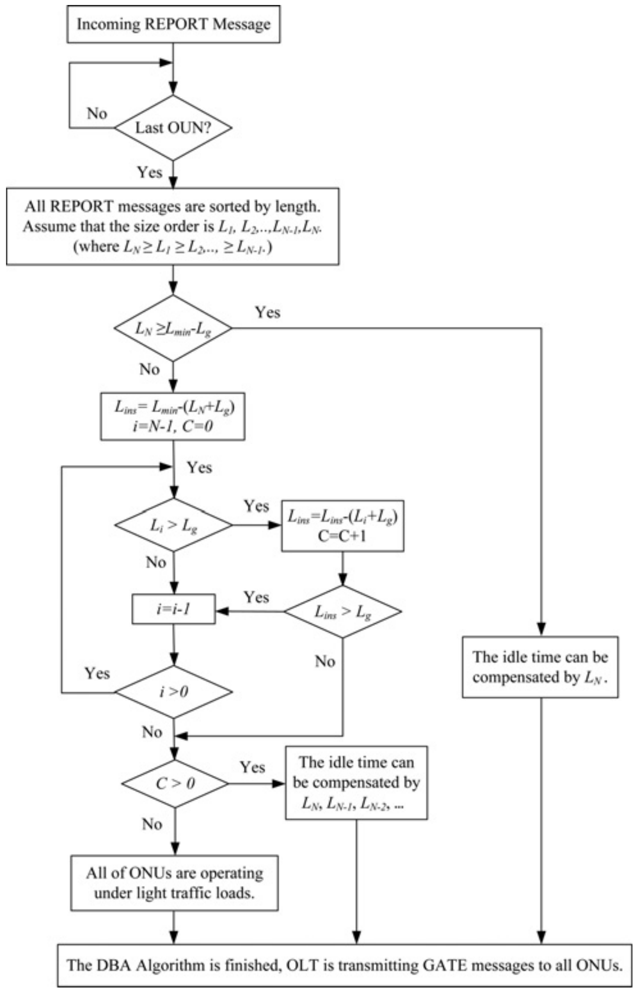

Figure 5 *Flowchart for the Sort-DBA algorithm*

(b) Packets arrive at ONU_i according to an independent, identically distributed (i.i.d.) Poisson process at the rate λ_i. Hence, the aggregate arrival rate of all ONUs is $\lambda = \sum_{i=1}^{N} \lambda_i$.

(c) Each ONU generates packets at the same average rate and each sends an equal amount of traffic to the OLT (uniform and symmetric traffic).

(d) Each ONU generates IP packets with a size distribution that matches the trace observed in an MCI backbone with OC-3 links (Fig. 6) [14]. The mean packet size for this distribution can be calculated as 353.8 bytes.

To validate our theoretical analysis, we simulate the performance of an EPON running the Sort-DBA algorithm. It is important to note that all simulations are run for sufficient time to obtain steady-state results. In general, one billion (1 000 000 000) time units are simulated per point in each curve. Simulations are conducted by SIMSCRIPT II code, and each experimental value is calculated by the variance reduction technique with 40 replicated simulations using different seeds. The results are obtained with 95% confidence.

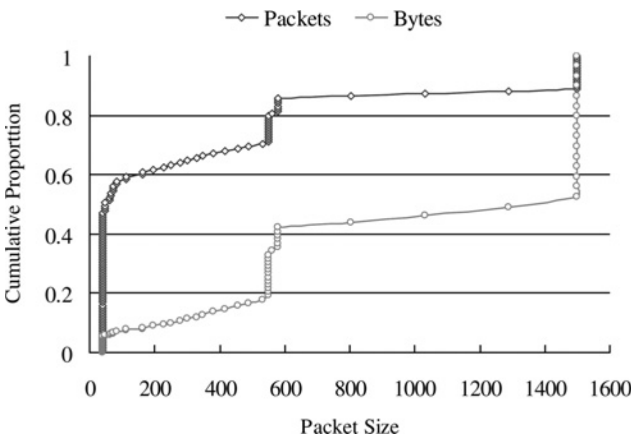

Figure 6 *OC-3 traffic distribution used in the simulations*

We also simulated the performance of four other DBA algorithms for comparison: IPACT (fixed), IPACT (limited), DBA2 and SC-DBA in the same conditions. The full set of simulation parameters is specified in Table 1. In this simulation, we have $L_{\text{min}} = 24.375$ kB by formula (2). Assuming that the guard-band time $T_g = 1\ \mu\text{s}$, the maximum transmission cycle length is $T_{\text{max}} = 3.144$ ms by formula (6). The traffic loads on the 16 ONU are almost identical and the ONU buffer size is 10 Mbytes. Each ONU receives aggregated OC-3 traffic from the users at a maximum rate of 100 Mbps. The upstream bandwidth (R_{u}) from all ONU to the OLT is 1 Gbps. In the star architecture, RTT is the same for both downstream and upstream delays (200 µs for a 20 km trunk).

4.2 Simulation results

We now compare the average transfer delay of the five algorithms shown in Fig. 7. The Sort-DBA algorithm has the lowest ONU delay, especially at medium or heavy traffic load. This improvement is mainly because of the fact that the last ONU REPORT is transmitted before the last of the grant data. This allows the OLT to finish its DBA calculation early and broadcast it to all ONUs during data transmission, provided that the last ONU transmission length is long enough to use up the idle time between cycle times. Moreover, the queue management method also decreases transfer delay because much of the REPORT overhead is not transmitted when ONU_i is operating under heavy load conditions. In contrast, the IPACT (fixed) algorithm has the worst performance because of the wasted idle periods and idle time slots. The IPACT (limited) algorithm, which dynamically allocates upstream bandwidth, performs somewhat better but still wastes bandwidth because of the wasted grant timeslots and the REPORT messages. DBA2 fares are even better, as it considers all REPORT messages and then allocates the idle slots of lightly loaded ONU to meet the bandwidth demand of the other heavily loaded ONU. Although it achieves higher utilisation, its advantages are limited because the early allocation mechanism cannot eliminate

doi: 10.1049/iet-com.2009.0587

Table 1 Simulation parameters for DBA algorithms in an EPON

	IPACT (fixed)	IPACT (limited)	DBA2	SC-DBA	Sort-DBA
numbers of ONU (N)	16	16	16	16	16
upstream bandwidth (R_u), Gbps	1	1	1	1	1
distance between OLT and ONUs, km	20	20	20	20	20
light velocity in fibre (V), km/s	2×10^5	2×10^5	2×10^5	2×10^5	2×10^5
queue discipline	FIFO	FIFO	FIFO	FIFO	FIFO
buffer size of ONU, Mbytes	10	10	10	10	10
L_{min}, kbytes	15	15	15	15	24.375
L_{min}-folds (X)	X	X	X	X	5
report length (H), bytes	64	64	64	64	64
guard time (T_g), μs	1	1	1	1	1
DBA processing time (T_{dba}), μs	0				0, 10, 100, 1000
ONU processing time, T_{ONU}	neglected				

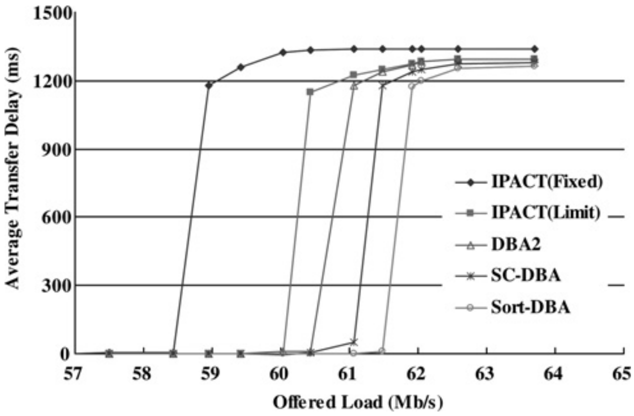

Figure 7 *Average transfer delay per ONU for IPACT (fixed), IPACT (limited), DBA2, SC-DBA and Sort-DBA algorithms*

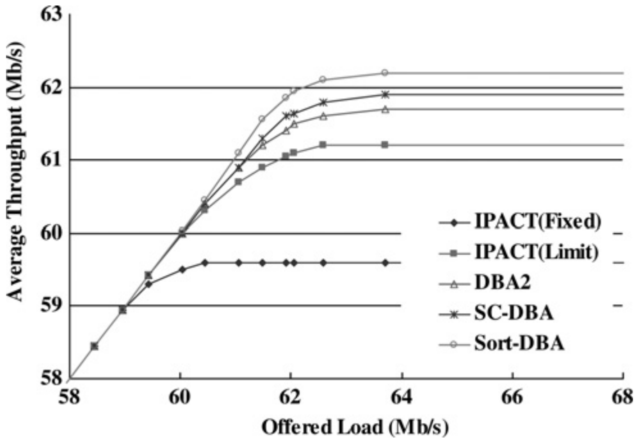

Figure 8 *Average throughput per ONU for IPACT (fixed), IPACT (limited), DBA2, SC-DBA and Sort-DBA algorithms*

the idle time when the previous ONUs are operating under heavy load conditions. SC-DBA improves upon the limitations of DBA2, but the achievement is limited since it cannot reduce the overhead of the REPORT messages when the ONU has a long queue length.

In Fig. 8, the overall throughput per ONU is similar for all DBA algorithms until the offered level reaches a value of about 59 Mbps. By formula (9), considering enhanced bandwidth utilisation by the queue management, the maximum throughput per ONU for Sort-DBA is 62.15 Mbps for $T_g = 1$ μs and $H = 64$ bytes. This result conforms to the simulation results, which clearly shows the excellent performance of the Sort-DBA algorithm. The Sort-DBA algorithm saves a significant amount of bandwidth by reducing the overhead of the REPORT messages under heavy load conditions. By transmitting the last of the REPORT messages before the last of grant data to reduce idle time, the proposed DBA is also able to serve needy ONUs earlier than IPACT (fixed), IPACT (limited), DBA2 or SC-DBA algorithms.

Fig. 9 shows the average queue length as a function of the offered load. Again, the Sort-DBA algorithm shows the best performance. Its queue length is the shortest for traffic loads between 55 and 62 Mbps, where the offered traffic pushes the limits of the network. Fig. 10 shows the average grant length as a function of the offered load. At medium and heavy loads, the Sort-DBA algorithm increases the demand length of each ONU to L_{min} to eliminate idle time, so the average grant length per ONU is more than the other DBA algorithms. The IPACT (fixed) algorithm defines a maximum transmission window of 15 kbytes per ONU in a

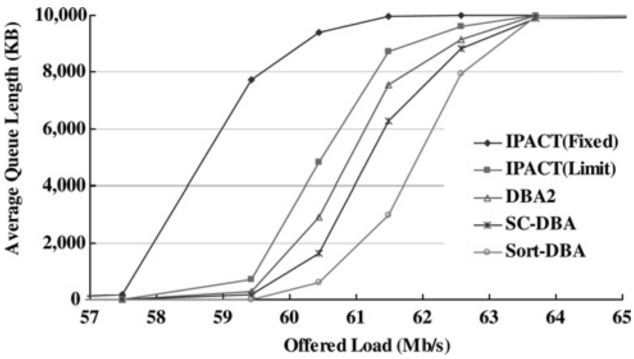

Figure 9 *Average queue length per ONU for IPACT (fixed), IPACT (limited), DBA2, SC-DBA and Sort-DBA algorithms*

cycle time of 2 ms, so its average grant length per ONU under heavy load is only 14.5 kbytes in our traffic model. The reason is fragmented Ethernet frames are not allowed, so it is unable to fully utilise the maximum transmission window with the FIFO queue discipline. The IPACT (limited) algorithm has the same result, but its average grant length is smaller than the IPACT (fixed) algorithm from light to high loads because IPACT (limited) always grants the requested data with the real load demand of each ONU so as to obtain the shortest transmission cycle. The DBA2 and SC-DBA algorithms always allocate the remaining idle slots to other ONUs when the transmission demand is excessive, allowing the average grant length to reach 15 kbytes.

Figs. 11 and 12 illustrate the average transfer delay and grant length with the Sort-DBA algorithm as a function of the offered load, with consideration of the impact of DBA calculation time (T_{dba}). Fig. 11 shows that the average transfer delay is almost unaffected by DBA calculation time because the proposed Sort-DBA algorithm considers T_{dba} as a factor of $L_{\min}$. The impact of T_{dba} affects only the average grant length shown in Fig. 12. The reason is that by this method, the computation of $L_{\min}$ automatically includes the T_{dba} factor. However, to implement this scheme, the OLT should notify the ONUs whether to send REPORT before data or send data before REPORT. However, the ONUs must check the sending order. This will increase complexity at both the OLT and ONU sides.

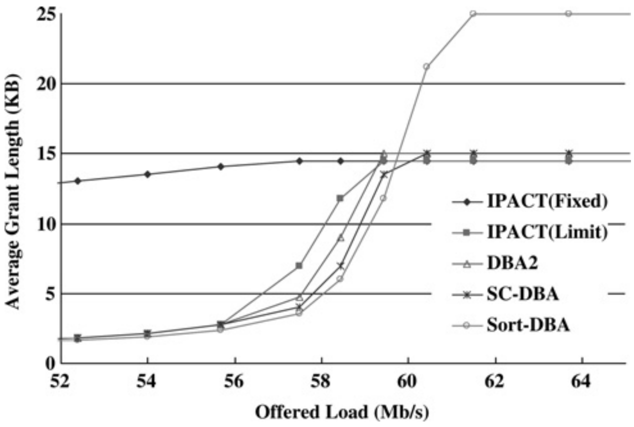

Figure 10 *Average grant length per ONU for IPACT (fixed), IPACT (limited), DBA2, SC-DBA and Sort-DBA algorithms*

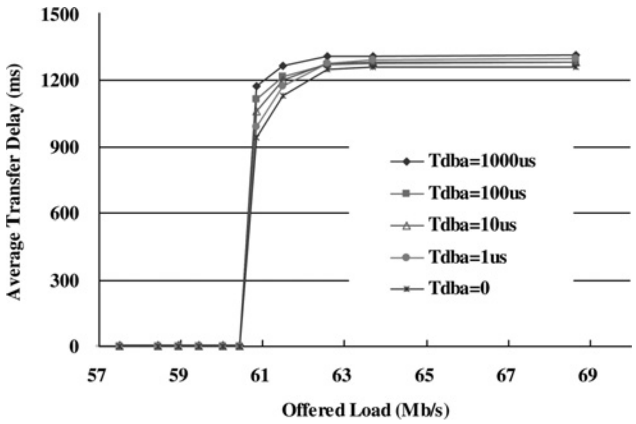

Figure 11 *Average transmission delay per ONU for Sort-DBA algorithm with consideration of DBA calculating time*

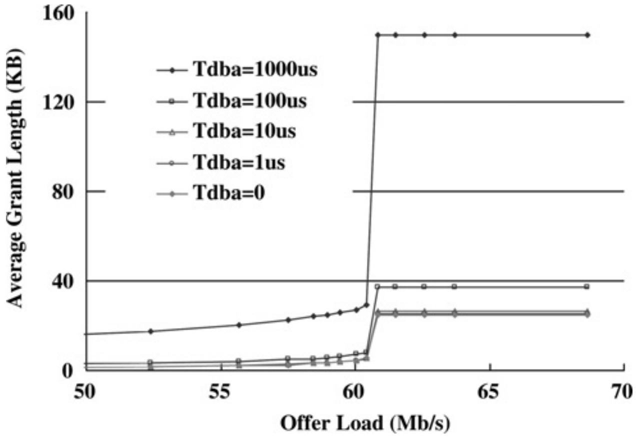

Figure 12 *Average grant length per ONU for Sort-DBA algorithm with consideration of the DBA calculating time*

The above discussion makes it clear that the Sort-DBA algorithm has the best performance with significantly more throughput per ONU than IPACT (fixed), IPACT (limited), DBA2 and SC-DBA under the same conditions. The proposed Sort-DBA algorithm also exhibits excellent bandwidth utilisation and low delay, making it suitable for EPON access networks.

5 Conclusions and future work

A Sort-based DBA algorithm has been proposed that sorts all REPORT messages with the request length so as to reduce the idle period and enhance bandwidth utilisation on EPONs. The proposed Sort-DBA provides collision-free upstream transmission and achieves higher utilisation

doi: 10.1049/iet-com.2009.0587

under most load conditions than other DBA algorithms. To fully utilise the idle time between transmission cycles, it calculates the maximum transmission window (unlike IPACT) for each ONU demand and adjusts the order of ONUs so that the longest transmission comes last. At medium and light loads, when none of the ONU transmissions is long enough to completely use up the idle time on its own, a compensation scheme is developed so that one or many grant data are laid out together after last grant data to fill the idle space. The presented simulation results clearly show that Sort-DBA allows significant improvement in bandwidth utilisation by reduction of the idle period. Furthermore, the proposed algorithm further enhances performance by using a queue management scheme to reduce the overhead of REPORT messages when the ONU queues are long. IPACT algorithms (with fixed service and limited service), DBA2 (an improved IPACT scheme) and the scheduling control scheme (an improved DBA2 scheme) were compared to our DBA algorithm in an event-driven EPON network simulation. The proposed DBA algorithm significantly achieves better performance than other DBA algorithms, with average maximum throughput per ONU and system efficiency of around 62.15 Mbps and 99.44%, respectively.

Our future work will investigate how to support a QoS scheme in a Sort-DBA algorithm. Secondly, we are extending our simulations to cases involving asymmetric traffic loads.

6 Acknowledgment

The authors would like to thank The National Science Council (NSC) of Taiwan for supporting this research under project number NSC 98-2221-E-151-037.

7 References

[1] KIM K.S.: 'On the evolution of PON-based FTTH solutions', *Inf. Sci.*, 2003, **149**, (1–3), pp. 21–30

[2] LUNG B.: 'PON architecture "future proofs" FTTH', *Lightw. Mag.*, 1999, **16**, (10), pp. 104–107

[3] ANGELOPOULOS J.D., BOUKIS G.C., VENIERIS I.S.: 'Delay priorities enhance utilisation of ATM PON access systems', *Comput. Commun.*, 1997, **20**, (11), pp. 937–949

[4] ANGELOPOULOS J.D., LELIGOU H.C., ARGYRIOU T., ZONTOS S., RINGOOT E., VAN CAENEGEM T.: 'Efficient transport of packets with QoS in an FSAN-aligned GPON', *IEEE Commun. Mag.*, 2004, **42**, (2), pp. 92–98

[5] MAIER G., MARTINELL M., PATTAVINA A., SALVADORI E.: 'Design and cost performance of the multistage WDM-PON access networks', *IEEE Lightw. Technol.*, 2000, **18**, (2), pp. 125–143

[6] SKUBIC B., CHEN J., AHMED J., WOSINSKA L., MUKHERJEE B.: 'A comparison of dynamic bandwidth allocation for EPON, GPON, and next-generation TDM PON', *IEEE Commun. Mag.*, 2009, **47**, (3), pp. S40–S48

[7] KRAMER G., MUKHERJEE B., PESAVENTO G.: 'Ethernet PON (ePON): design and analysis of an optical access network', *Photonic Netw. Commun.*, 2001, **3**, (3), pp. 307–319

[8] CHAE C.J., WONG E., TUCKER R.S.: 'Optical CSMA/CD media access scheme for ethernet over passive optical network', *IEEE Photonics Technol. Lett.*, 2002, **14**, (5), pp. 711–713

[9] KRAMER G., PESAVENTO G.: 'Ethernet passive optical network (EPON): building a next-generation optical access network', *IEEE Commun. Mag.*, 2002, **40**, (2), pp. 66–73

[10] KRAMER G., MUKHERJEE B., PESAVENTO G.: 'IPACT: a dynamic protocol for an Ethernet PON (EPON)', *IEEE Commun. Mag.*, 2002, **4**, (1), pp. 74–80

[11] LUO Y., ANSARI N.: 'Limited sharing with traffic prediction for dynamic bandwidth allocation and QoS provisioning over EPONs', *J. Opt. Netw.*, 2005, **4**, (9), pp. 561–572

[12] ASSI C., YE Y., DIXIT S., ALI M.A.: 'Dynamic bandwidth allocation for quality of service over Ethernet PONs', *IEEE J. Sel. Area. Commun.*, 2003, **21**, (9), pp. 1467–1477

[13] ZHENG J.: 'Efficient bandwidth allocation algorithm for Ethernet passive optical networks', *IEE Proc. Commun.*, 2006, **153**, (3), pp. 464–468

[14] IEEE 802.3ah task force, http://www.ieee802.org/3/efm

[15] WAN packet size distribution available from http://www.nlanr.net/NA/Learn/packetsizes.html

[16] ZHU Y., MA M.: 'IPACT with grant estimation (IPACT-GE) scheme for Ethernet passive optical networks', *J. Lightw. Technol.*, 2008, **26**, (14), pp. 2055–2063

doi: 10.1049/iet-com.2009.0587