Ring-based local access PON architecture for supporting private networking capability

ASM Delowar Hossain, R. Dorsinville, and M. A. Ali

Department of Electrical Engineering, Graduate School of City University of New York, New York
ASMDHossain@gc.cuny.edu

A. Shami

Department of Electrical and Computer Engineering, University of Western Ontario, Ontario, Canada

C. Assi

Department of Engineering and Computer Science, Concordia University, Montreal, Canada

RECEIVED 7 SEPTEMBER 2005; REVISED 23 NOVEMBER 2005;
ACCEPTED 28 NOVEMBER 2005; PUBLISHED 04 JANUARY 2006

We propose, to our knowledge, a novel ring-based local access Passive Optical Network (PON) architecture that addresses some of the limitations of current tree-based PON architectures that include supporting private networking capability. Specifically, we propose and devise a simple ring-based Ethernet PON (EPON) architecture that supports a truly shared LAN capability among end users as well as upstream access to the central office. Unlike a typical ring-based PON topology in which the optical line terminal (OLT) and the optical network units (ONUs) are interconnected via a long fiber ring, under the proposed architecture, ONUs are interconnected via a short distribution fiber ring in the local loop but share the standard trunk feeder fiber for long-reach connectivity to the central office (CO). The main characteristic of the proposed architecture is that it supports a fully distributed control plane among the ONUs for ONU–ONU communication as well as upstream access to the OLT. This architecture is well suited to an autonomous access environment such as a university campus or a private corporation where several buildings are closely dispersed within a 0.5–1 km diameter area. © 2006 Optical Society of America

OCIS codes: 060.2330, 060.4250.

1. Introduction

Recent advances in optical networking technologies have fueled tremendous growth in both backbone and Metropolitan Access Network (MAN) capacity. At the same time, the performance of end users' computing equipment reached gigahertz speeds. The conduits linking the high-speed end-user equipment to the high-capacity backbone networks, however, remain a bottleneck. These connections are commonly referred to as the "first mile." While recent advances in "first mile" technology have increased capacity from the range of 56 kb/s for a dial-up modem to a few megabytes per second for a cable modem (CM) or digital subscriber line (DSL) connection, this is still far short of the gigabit line speed necessary to support rich multimedia and real-time services.

Fiber-to-the-home (FTTH) is the ultimate level of access, allowing end users to access the backbone networks through the gigabit capacity of a fiber optic cable. Unfortunately, current systems have proven too complex and expensive to be commercially viable. To lower the cost and expedite the implementation of FTTH, Passive Optical Network (PON) based-solutions have been proposed [1]. It is likely that the reduced equipment costs and the reduced operational costs of PONs will enable carriers to justify FTTH, thus solving the "first mile" bottleneck.

Ethernet-based PON (EPON) technology is emerging as a viable choice for the next-generation broadband access network [2–4]. A PON is a point-to-multipoint fiber optical network with no active elements in the signal's path. Each customer receives a dedicated short optical fiber but shares the long distribution trunk fiber. All transmissions in a PON are performed between an optical line terminal (OLT) and optical network units (ONUs) using typical Ethernet frames. Traffic from an OLT to an ONU is called "downstream" (point-to-multipoint), and traffic from an ONU to the OLT is called "upstream" (multipoint-to-point). Two wavelengths are used: typically $1310\,\text{nm}(\lambda_{\text{up}})$ for the upstream transmission and $1490\,\text{nm}(\lambda_d)$ for the downstream transmission. In the downstream direction, an EPON operates as a broadcast and select network. In the upstream direction, multiple ONUs share the transmission channel. Thus, the ONUs need to employ some arbitration mechanism to avoid collisions. To date, PONs and almost all proposed EPON architectures reported in the literature are deployed in tree-and-branch configurations.

In general, the OLT arbitrates upstream transmissions by allocating an appropriate time slot to each ONU. Only one ONU is allowed to transmit during a designated time slot. The start time and the duration of such a time slot is calculated via a dynamic bandwidth allocation (DBA) module, residing at the OLT. To facilitate the implementation of bandwidth allocation schemes, the OLT and ONUs exchange control messages, namely, REPORT message and GATE message. These control messages are defined by the IEEE 802.3ah task force through the development of Multi-Point Control Protocol (MPCP) [5]. A REPORT message is sent by an ONU to OLT informing it about its bandwidth requirements. Upon receiving a REPORT, the OLT passes the message to a DBA module to perform the bandwidth allocation computation. The OLT then grants the ONU a transmission time slot by sending a GATE message indicating the start time and the duration of such a time slot.

Deploying point-to-point customer communication links to emulate a shared local area network (LAN) within a single PON infrastructure is an important feature for providing a private networking capability, which has recently received some attention [6–8]. In general, standard upper-layer shared LAN emulation techniques require the use of bridges/routers at the OLT to redirect data frames back to the ONUs [6]. These techniques are effective but reduce the available downstream transmission bandwidth and increase end-to-end delay of LAN traffic. Several physical layer LAN emulation techniques have been proposed to achieve intercommunication among ONUs within the same PON setup [7, 8]. These include a physical layer solution that employs a fiber Bragg grating at the trunk feeder fiber to reflect a predefined wavelength channel for intercommunication between ONUs [7]. Another approach assumes that each ONU is connected to a star coupler (SC) via two distribution fibers, where one fiber is used to transport both downstream and upstream traffic and the second fiber is used to deliver redirected frames back to the ONU for detection [8]. Most of these solutions suffer from high splitting loss as the redirected signal traverses through the SC once or twice, resulting in a lower power budget that limits the number of ONUs that can be attached to a single PON.

To provide a simple PON-based local-access infrastructure that addresses these limitations and supports private networking capability, this work proposes a novel ring-based EPON architecture that supports a truly shared LAN capability as well as upstream access to the OLT (MAN/WAN traffic). Unlike a typical ring-based PON topology in which OLT and ONUs are interconnected via a long fiber ring, under the proposed architecture, ONUs are interconnected via a short distribution fiber ring in the local loop but share the standard trunk feeder fiber for long-reach connectivity to the OLT. This minimizes fiber deployment in both the CO and the local loop. This architecture is well suited to an autonomous access environment such as a university campus or a private corporation where several buildings are closely dispersed within a 0.5–1 km diameter area.

The main characteristic of the proposed architecture is that it supports a fully distributed control plane among the ONUs for ONU–ONU communication as well as upstream access to the OLT. Specifically, it utilizes a fully distributed time division multiple access (TDMA) arbitration scheme in which the OLT is excluded from the arbitration process. In the proposed decentralized scheme, the ONUs exchange signaling and control information concerning their queue status and their transmission needs among themselves (REPORT messages). Then, the ONUs concurrently and independently run instances of the same DBA algorithm, outputting identical bandwidth allocation results. Once the algorithm is run, the ONUs sequentially and in orderly fashion transmit their data including both LAN and MAN/WAN traffic without any collisions, eliminating the OLT's centralized task of processing requests and generating grants for bandwidth allocations.

In addition to the added flexibility and reliability of a distributed scheme, the proposed ring-based PON architecture introduces the following several advantages over a typical star-based centralized PON architecture: (1) supports a truly shared LAN capability within a single PON setup; (2) eliminates the typical utilization of an upstream burst-mode receiver and associated design challenges at the OLT; (3) increases the available downstream and upstream channel bandwidth; (4) supports more efficient upstream channel utilization; and (5) alleviates the typical limited power budget problem associated with physical layer LAN emulation techniques described above.

The rest of the paper is organized as follows: Section 2 describes the proposed distributed architecture. Section 3 discusses the distributed bandwidth allocation algorithm. Section 4 provides performance results for the proposed scheme, and Section 5 offers some concluding remarks.

2. Proposed Architecture

Figure 1(a) illustrates the proposed ring-based PON architecture. An OLT is connected to N number of ONUs via a 20 km trunk feeder fiber, a passive 3-port optical circulator, and a short distribution fiber ring. To cover the same local access area as that covered by a conventional tree-based architecture [3, 4], the small ring at the end of the trunk is assumed to have a 1 km diameter. The set of ONUs are joined by point-to-point links in a closed loop. The links are unidirectional: both downstream and upstream signals (combined signal) are transmitted in one direction only. Figure 1(b) shows detailed ONU architecture. Each ONU attaches to the ring at a $(n:\ 1-n)$ 1×2 passive star coupler [incoming signal at point A in Fig. 1(b)] and can transmit data onto the ring through the output port of a 2×1 passive combiner [outgoing signal at point C in Fig. 1(b)]. Note that in addition to the conventional transceiver maintained at each ONU (a λ_{up} upstream transmitter and a λ_d downstream receiver), this approach requires an extra receiver tuned at λ_{up}.

The downstream signal is coupled to the ring at port 2 of the optical circulator. After recombining with the recirculated upstream signal via a 2×1 passive coupler [Fig. 1(a)] placed on the ring directly after the optical circulator, the combined signal then circulates around the ring (ONU 1 through ONU N) in a Drop-and-Go fashion. The downstream signal is then removed at the end of the ring using a filter (located directly after the last ONU) that passes only the 1310 nm upstream signal. The upstream signal emerging from the filter at the end of the ring is split into two components via a 1×2 passive splitter [Fig. 1(a)] placed on the ring directly after the filter. The first component is directed toward the OLT via circulator ports 1 and 3, while the second component is allowed to recirculate around the ring after recombining with the downstream signal (originating from the OLT) via the 2×1 coupler of Fig. 1(a).

The $(n$: 1-$n)$ 1×2 coupler (n is a small arbitrary percentage assumed here to be 10%) splits the incoming combined signal at each node into a small (10%) "Drop-signal portion"

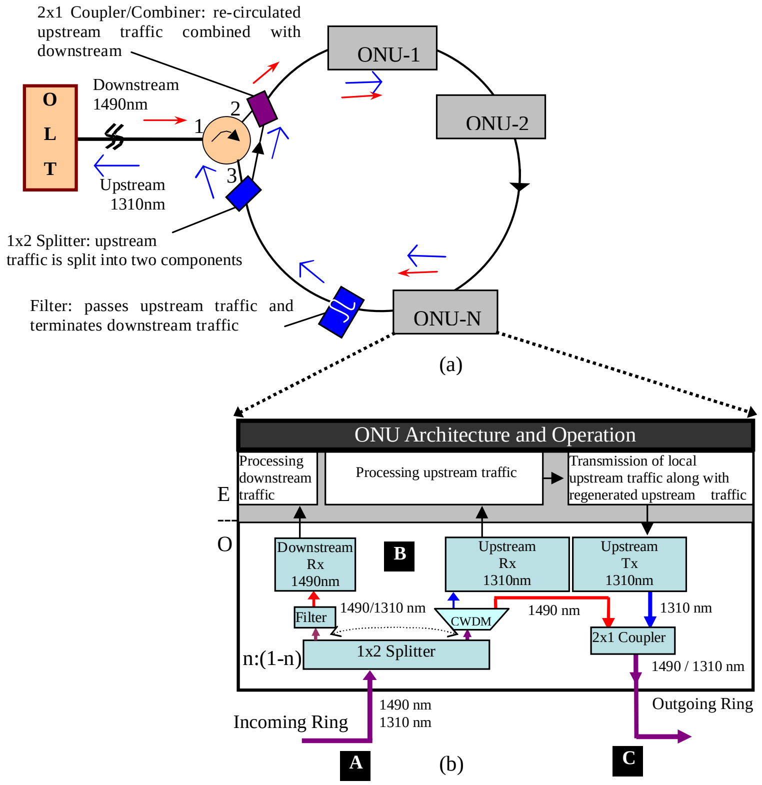

Fig. 1. (a) Proposed ring-based architecture, (b) ONU architecture.

and a large (90%) "Go-signal portion." The small portion of the circulating combined signal dropped at each node (Drop-signal) is passed through a filter that removes the upstream signal and passes only the downstream broadcast signal, which is then received and processed by the 1490 nm downstream receiver. The remaining portion of the combined signal emerging from the 90% coupler's port (Go-signal) is first separated into its two constituents: downstream and upstream signals via a coarse wave division multiplexing (CWDM) filter. The separated upstream signal (second component) is received and processed via the 1310 nm upstream optical receiver housed at the ONU, where it is then regenerated and retransmitted along with the ONU's own local control and data traffic. Finally, the separated downstream signal is recombined with the retransmitted upstream signal (regenerated plus local) via the 2 × 1 coupler of Fig. 1(b) to form the outgoing combined signal (incoming combined signal for next ONU) that circulates around the ring.

Since upstream transmission is based on a TDMA scheme, inter-ONU traffic (LAN data and control messages exchanged among ONUs) is transmitted along with upstream traffic destined to the OLT (MAN/WAN data) within the same preassigned time slot. The first component of the upstream signal destined to the OLT is received and processed by the 1310 nm upstream optical receiver (housed at the OLT), which accepts only MAN/WAN traffic; discards LAN traffic; and may discard or process (for reasons to be given below) the control messages. On the other hand, the second component of upstream signal is transmitted sequentially, bit by bit, around the ring from one node to the next, where it is regenerated and retransmitted at each node.

Since the ring is a closed loop, upstream traffic will circulate indefinitely unless removed. The process of removing, regenerating, and retransmitting the second component of the upstream signal at each node (ONU) is implemented as follows. First, the 1310 nm upstream optical receiver (housed at each ONU) terminates all upstream traffic, examines the destination media access control (MAC) address of each detected Ethernet frame, and then performs one or more of the following functions: (1) all recirculated upstream traffic addressed to the OLT is removed by the first ONU [ONU that is physically located on the ring directly after the 2 × 1 coupler of Fig. 1(a)]; (2) all control messages (REPORTs) must be processed, regenerated, and then retransmitted by each node; (3) the source node removes its own transmitted inter-ONU control messages that complete one trip around the ring through recirculation; (4) transient LAN traffic, terminated at an intermediate node but destined to other nodes are regenerated and then retransmitted along with the node's own local upstream traffic within the designated proper time slot; (5) once the destination address of the LAN traffic matches the node's MAC address, it is copied and delivered to the end users and then discarded (not retransmitted to the next ONU). Note that only inter-ONU control traffic completes one trip around the ring (returns to the source node); the LAN traffic is removed at the destination ONU, and the MAN/WAN traffic destined to the OLT is always removed before returning to the source node except for first ONU's traffic, which is removed by the source node (first ONU).

Note that any node (ONU) failure may bring down the whole network. Thus, protecting both node and fiber failures is essential for proper operation of the proposed network. For instance, to ensure failure recovery, there could be redundancy of the ONU transceiver and fiber. However, failure recovery is beyond the scope of this work and will be addressed in another paper.

2.A. Downstream and Upstream Typical Limited Optical-Power-Budget Problem

Since the upstream signal is regenerated at every node, the typical limited-power-budget problem (due to the splitting loss at each node) as well as the receiver-dynamic-range problem (long/short optical network paths and different splitting factors) are totally eliminated. To examine the effect on performance of the downstream power budget, we consider the best- and worst-case scenarios by calculating the total optical distribution network (ODN) loss (passive optical elements such as splitters, combiners, fibers, connectors, and splices forming an optical path) incurred by the downstream signal on its shortest and longest optical paths from the OLT to the first and the last ONU, respectively.

We assume that the trunk feeder fiber length is 20 km, the first ONU is 20 km away from the OLT, and the last ONU is 23.2 km away from the OLT (ring circumference is approximately 3.2 km). We consider losses of components based on recent industry standards: circulator insertion loss is 0.3 dB, and trunk SMF fiber loss at 1490 nm is 0.2 dB/km. There are two types of losses encountered by the downstream signal at each node. The first type is from point A to B in Fig. 1(b) (Drop-component) and the second type is from point A to C in Fig. 1(b) (Go-component). Table 1 quantifies both types of losses. Thus, total ODN loss incurred by downstream signal on its path to the first and the last ONU is

$$L_{\text{total\_}Loss}^{\text{1st\_}ONU} = L_{\text{trunk}}^{\text{fiber}} + L_{\text{trunk}}^{\text{circulator}} + L_{A\text{-}B}^{\text{ONU}} = 14.6\,\text{dB}, \tag{1}$$

$$L_{\text{total\_}Loss}^{\text{last\_}ONU} = L_{\text{trunk}}^{\text{fiber}} + L_{\text{trunk}}^{\text{circulator}} + (N-1)\,L_{A\text{-}C}^{\text{ONU}} + L_{\text{ring}}^{\text{fiber}} + L_{A\text{-}B}^{\text{ONU}}. \tag{2}$$

For example, if $N = 16$ ONUs, the total ODN loss using Eq. (2) above (worst case) is 30.9 dB, leading to a 16.3 dB of ODN differential attenuation. Assuming that the signal power transmitted by the OLT is 5 dBm, a receiver sensitivity of -25.9 dBm will be required at the last ONU. Thus, a downstream receiver with a -25.9 dBm minimum sensitivity combined with a 16.3 dB dynamic range is required for support of 16 ONUs. Note that downstream power-budget demands on receiver parameters are still within practical and commercial reach.

Table 1. Typical Losses Incurred at an ONU

Type of Loss	Point A to B	Point A to C
Connector	0.15	2×0.15
Splitter (10/90)	10.0	0.45
Filter	0.15	—
CWDM	—	0.3
Total (dB)	10.3	1.05

Note also that the 16-node limit is not a shortcoming of an architecture that is specifically devised to support a private ring-based local access infrastructure within a 1 km diameter area. A typical large private organization would have at most 10–15 buildings within such a geographically bounded area. To scale beyond 16 ONUs, a much higher transmitted power, a highly sensitive receiver, and lower component losses are required. Alternative complex solutions to support large number of ONUs by virtually eliminating the downstream power budget problem can be achieved by regeneration of downstream signal at ONUs; but that will require additional downstream transmitter at ONUs, slightly increasing the ONU complexity and cost.

3. Distributed DBA Scheme

3.A. Overview of Centralized DBA Schemes

Several centralized tree-based DBA schemes have recently been reported in the literature [9–13]. An OLT-based polling scheme, called Interleaved Polling with Adaptive Cycle Time (IPACT) based on Grant and Request messages, has been presented in Ref. [9]. Using IPACT, several DBA schemes were studied in Ref. [9], namely, fixed, limited, gated, constant credit, and linear credit. Among these algorithms, the limited was shown to exhibit the best performance. The limited DBA scheme is cycle based, where a cycle (T_{CYC}) is defined as the time that elapses between two executions of the scheduling algorithm. A cycle has a variable length size confined within certain lower and upper bounds, which we denote as T_{MIN} and T_{MAX} (s) respectively. Thus, the algorithm schedules between B_{MIN} and B_{MAX} (bytes) at a time, where B_i is determined by multiplying T_i with the line rate. In this scheme, the ONU will be granted the requested number of bytes but no more than a given predetermined maximum B_{MAX}. If R_i is the requested bandwidth of ONU$_i$, then the granted bandwidth (B_{Granted}) is equal to

$$B_{\mathrm{Granted}} = \begin{cases} R_i & \text{if } R_i \leq B_{\mathrm{MAX}} \\ B_{\mathrm{MAX}} & \text{if } R_i > B_{\mathrm{MAX}} \end{cases} . \tag{3}$$

B_{MAX} is determined by the maximum cycle time T_{MAX} [9]:

$$B_{\mathrm{MAX}} = \frac{1}{N}\left[R_{\mathrm{EPON}}\left(T_{\mathrm{MAX}}\text{-}\left(N^{*}T_G\right)\right)\right],$$

where N is the number of ONUs, T_G is the guard band slot, and R_{EPON} is EPON line rate.

All of the above-referenced DBA schemes are OLT based; that is, the OLT has centralized intelligence. The performance of most of these centralized schemes, including the limited scheme, suffers from several limitations: (1) The bandwidth granted by the OLT, during cycle n, to ONU$_i$ is determined only by the content of a single REPORT message transmitted in the previous cycle $n - 1$ by ONU$_i$ (i.e., the bandwidth computation module does not take into account the remaining requests of other ONUs). Thus, the process of bandwidth allocation is not globally optimized; (2) due to the bursty nature of Ethernet traffic, some ONUs might have less traffic to transmit while other ONUs may require more bandwidth than B_{MAX}. For instance, assume that ONU$_i$ requests an amount of bandwidth $R_i < B_{\mathrm{MAX}}$, while ONU$_j$ requests an amount of bandwidth $R_j > B_{\mathrm{MAX}}$. Although there is an excess amount of bandwidth $(B_{\mathrm{MAX}} - R_i)$ that can be granted to ONU$_j$, due to limitation #1 cited above, the maximum bandwidth that may be granted to ONU$_j$ is only B_{MAX}.

3.B. Proposed DBA Scheme under a Distributed Control Plane

The proposed scheme assumes a cycle-based upstream link, where the cycle size can be either fixed, or variable length confined within certain lower and upper bounds to accommodate the dynamic upstream traffic conditions. During a given cycle, each ONU transmits its control (REPORT) message (within its assigned time slot) around the ring from one node to the next, where it is finally removed by the source ONU after making one trip around the ring. It typically contains the desired size of the next time slot based on the current ONU's buffer occupancy. Since these message frames are typical Ethernet frames, the ONU should also account for additional Ethernet overhead when requesting the next time slot; this includes an 8-byte frame preamble and a 12-byte Inter Frame Gap (IFG) associated with each frame. Note that since this overhead is typical of the Ethernet frame, there is no additional overhead due to this architecture. Thus, this framing has no effect on our network besides what is standard with any EPON.

Since the REPORT messages are processed and retransmitted at each node, ONUs can directly communicate their status and exchange signaling and control message information with one another. Because both REPORT message and data frames (LAN and MAN/WAN data) are transmitted within the same time slot, a REPORT message can be transmitted at either the beginning or the end of a time slot, depending on the bandwidth-request approach implemented by the ONU. For reasons to be explained below, this work assumes that control messages (REPORTs) are always transmitted at the beginning of a time slot, followed by LAN data, and finally followed by MAN/WAN data.

Each ONU maintains a database that contains the state of the queues of all the ONUs. This information is updated each cycle whenever the ONU receives new REPORT messages from all other ONUs. The DBA module housed at each ONU uses this information to calculate a new set of time slot assignments at each cycle. The ONUs sequentially and independently run instances of the same DBA algorithm, outputting identical bandwidth allocation results. The execution of the algorithm at each ONU starts immediately once all REPORT messages have been collected. Thus, all ONUs must execute the DBA algorithm prior to the expiration of the current cycle so that bandwidth allocations scheduled for the next cycle are guaranteed to be ready by the end of current cycle. Note that normally a cycle starts after the DBA calculation for ONU time slots is completed. Each cycle consists of N time slots, where N is the number of ONUs in the network. Once all ONUs of a cycle have completed their transmissions, then that cycle is over. An execution of the DBA algorithm produces a unique and identical set of ONU assignments. It is critical that the algorithm produce a unique outcome for any arbitrary set of inputs. Once the algorithm has been executed, the ONUs sequentially and in orderly fashion transmit their data without any collisions, eliminating the OLT's centralized task of processing requests and generating grants for bandwidth allocations.

To carry out a general and fair comparison between the performances of a centralized DBA scheme and the proposed distributed scheme, this work uses the centralized-limited-service scheme reported in Ref. [9], along with the appropriate changes needed to accommodate the distributed architecture, as the basis for the decentralized DBA scheme presented here. As mentioned above, to globally optimize the bandwidth allocation process, the proposed DBA algorithm execution is performed only after each ONU receives and processes all other ONUs requests.

Based on bandwidth demands, ONUs can be classified into two groups, namely, lightly loaded ONUs that have bandwidth demands less than B_{MAX}, and heavily loaded ONUs that have bandwidth demands more than B_{MAX}. During each cycle, the DBA module must keep track of the unclaimed bandwidth from the set of lightly loaded ONUs. It then must redistribute this excess bandwidth to other heavily loaded ONUs based on their requested bandwidths; i.e., two ONUs requesting bandwidths $B1$ and $B2$ more than B_{MAX} will be assigned excess bandwidths proportional to $B1$ and $B2$.

During each cycle, the lightly loaded ONUs with $R_i < B_{\mathrm{MAX}}$ will contribute a total remainder cycle bandwidth:

$$B_{\mathrm{CycleRemainder}} = \sum_{i}^{L} (B_{\mathrm{MAX}} - R_i) \quad L: \text{Number of lightly loaded ONUs.}$$

The heavily loaded ONUs with $R_i > B_{\mathrm{MAX}}$ will require a total over the limit cycle bandwidth:

$$B_{\mathrm{CycleOverLimit}} = \sum_{i}^{H} (R_i - B_{\mathrm{MAX}}) \quad H: \text{ Number of heavily loaded ONUs.}$$

The total remainder cycle bandwidth can be fairly distributed among the heavily loaded ONUs to expand their maximum transmission window as follows [10]:

$$\Delta B_i^{\mathrm{extra}} = B_{\mathrm{CycleRemainder}} \left[\frac{R_i - B_{\mathrm{MAX}}}{B_{\mathrm{CycleOverLimit}}} \right],$$

where ΔB_i is the extra bandwidth allocated to ONU_i. The granted bandwidth, B_{GH}, for a heavily loaded ONU_i is given by

$$B_{GH} = \Delta B_i^{\mathrm{extra}} + B_{\mathrm{MAX}}. \tag{4}$$

If R_i is the requested bandwidth of ONU$_i$, B_{Granted} is the bandwidth granted using the proposed limited service-based distributed DBA scheme [Eqs. (3), (4)], then B_{Granted} can be expressed as

$$B_{\text{Granted}} = \begin{cases} R_i & \text{if } R_i \le B_{\text{MAX}} \\ R_i & \text{if } R_i > B_{\text{MAX}} \text{and} B_{\text{CycleRemainder}} \ge B_{\text{CycleOverLimit}} \\ B_{GH} & \text{if } R_i > B_{\text{MAX}} \text{and} B_{\text{CycleRemainder}} < B_{\text{CycleOverLimit}}. \end{cases}$$

Note that the lightly loaded ONUs $(R_i < B_{\text{MAX}})$ can be scheduled instantaneously "on the fly" without waiting for DBA module to perform its end-of-cycle computations, whereas the heavily loaded ONUs $(R_i > B_{\text{MAX}})$ will have to wait until all REPORT messages have been received and the DBA algorithm has computed their bandwidth allocations. Thus, lightly loaded ONUs can be scheduled ahead of heavily loaded ones. The drawback of this approach is that it introduces some idle time between two consecutive cycles, n and $n+1$, provided that the ONUs follow the conventional approach [9–13] and transmit their REPORT messages after data frames. This idle time, during which the PON channel is not utilized, is equal to the time of one trip around the ring (to collect the REPORT) plus the time to perform the DBA computation.

To eliminate this idle time (as proposed above), for a given time slot, the REPORT message is always transmitted before data frames. Assume that the duration of the time slot allocated only to the data transmission of the last ONU (the ONU that is scheduled to transmit last) is always longer than the idle time. This assumption is always valid since an ONU that is scheduled last must always be a heavily loaded ONU. In this case, all other ONUs would have ample time to complete the execution of the DBA algorithm prior to the completion of the last ONU's data transmission (within the same cycle). This ensures that the process of bandwidth allocations to cycle $n+1$ is always executed before the expiration of cycle n.

Note that, in contrast to centralized architectures where the order of the ONU's transmission is fixed (sequential) in each cycle, the distributed architecture has the added flexibility of varying the order of the ONU's transmission according to the ONU's traffic demands and priority. Thus, the order of ONU's transmission may be different in each cycle and need not be fixed. As mentioned above, this work assumes that lightly loaded ONUs are always scheduled ahead of heavily loaded ones. If there is more than one lightly loaded ONU at a given cycle n, we assume that the ONUs' transmission order is based on the transmission order of the previous cycle $(n-1)$; that is, an ONU that reports its queue status first is scheduled first. We further assume that the order of a heavily loaded ONU's transmission is based on the shortest request first (SRF); that is, the ONU with the shortest bandwidth request is scheduled first (in ascending order), and ties are broken by the ONU ID.

Note that, if ONUs transmissions are scheduled in ascending order, upstream data transmissions assigned to any consecutive time slots are always collision free. In contrast, if ONUs' transmissions are scheduled in a descending order (for instance, ONU5 is scheduled to transmit first and then ONU2), collision between two consecutive time slots is possible (recirculated upstream data transmitted at the end of the time slot assigned to ONU5 might collide with data transmitted at the beginning of the next scheduled time slot assigned to ONU2). To avoid this problem, as suggested above, upstream MAN/WAN data are always transmitted at the end of a time slot. Since MAN/WAN traffic destined to the OLT is always removed by the first ONU before returning to the source node, i.e., it is not allowed to recirculate around the ring, collision-free transmission is ensured. Furthermore, since LAN traffic is terminated at the destination node and is not allowed to recirculate either (except when traffic is addressed to an ONU located opposite to the direction of traffic flow), this further ensures a collision-free transmission.

It is important to emphasize that while the present work considers only a distributed control plane among the ONUs to achieve direct intercommunication between them as well as upstream access to the OLT, a simple OLT-based centralized control plane can also be implemented along with the proposed distributed one, resulting in a hybrid control plane. In a system that supports a hybrid control plane, the OLT, rather than discarding upstream control messages, can process these messages to track and monitor both LAN and upstream MAN/WAN traffic levels. The OLT may issue a downstream control signal (GRANT message) to halt or prioritize certain traffic or services if an ONU violates its predetermined service level agreement.

3.C. *Synchronization*

All ONUs are synchronized to a common reference clock extracted from the OLT downstream traffic. Clocking information, in the form of a synchronization marker, is included at the beginning of each downstream frame cycle. The synchronization marker is a 1 byte code that is transmitted every 2 ms to synchronize the ONUs with the OLT [5]. The TDM controller at each ONU, in conjunction with timing information from the OLT, controls the upstream transmission of the variable-length packets within the dedicated time slots. Maintaining proper time synchronization between different ONUs is required for the appropriate operation of the distributed DBA algorithm.

4. Performance Evaluation

In this section, we assess the feasibility and compare the performance of the proposed architecture that uses the distributed DBA scheme with that of the centralized architecture that uses the limited DBA scheme. An event-driven packet-based simulation model was developed using C++. Two simulation programs with identical network parameters were developed, one for the centralized IPACT architecture and the other for the decentralized architecture. The performance metrics used here are average packet queuing delay and upstream channel utilization.

To compare the performance results of the proposed distributed scheme with that of the centralized scheme of [9], we use the same system parameters used therein: a system with 16 ONUs, access link data rate from users to an ONU of 100 Mb/s, and 1 Gb/s upstream EPON line rate (from an ONU to the OLT). The distance between the OLT and the ONUs varies from 20 to 23 km (ring circumference ≈ 3 km). Maximum cycle time T_{MAX} is 2 ms. For the centralized architecture, the guard time, T_G, separating two consecutive transmission windows is set to 5 µs, [9], whereas for the distributed architecture, we set $T_G = 0$ (since guard time slots are not needed). Buffer size in each ONU is 10 Mbytes.

The traffic model used here is the same as that reported in Ref. [9], where each ONU has a number of ON/OFF sources, each with a Pareto distribution governing the lengths of the ON/OFF periods, in order to capture the self-similar nature of Ethernet traffic [14, 15]. All arriving frames are then queued in a first-in-first-out buffer. Each point on the following plots corresponds to a sample of 50 million packets averaged over four different runs. In the following two subsections, the simulations will be repeated for two types of traffic. In Subsection 4.A, all upstream traffic is assumed to be destined to the OLT (MAN/WAN traffic). In Subsection 4.B, upstream traffic is assumed to be a mixture of MAN/WAN and shared LAN traffic.

4.A. *Performance of Typical Upstream MAN/WAN Traffic*

In this subsection, we assume that all upstream traffic is destined to the OLT (MAN/WAN traffic). Figures 2, 3 compare the average packet queuing delay and channel utilization as a function of offered ONU load (OOL) for both the centralized and the distributed DBA schemes. As can be seen from Figs. 2, 3, the distributed architecture demonstrates an improvement over the centralized approach in terms of average packet delay and channel utilization. At a very low load (0.1–0.2), average packet delay of the distributed DBA scheme is approximately $120-180\,\mu\text{s}$ less than that of the centralized scheme (Fig. 2). This is due mainly to the savings of guard band slots ($5\,\mu\text{s}/\text{ONU}^{*}16$ ONUs $= 80\,\mu\text{s}$) as well as to interchanging the order of ONUs transmissions, since queuing (TDM) delay is minimal for both architectures.

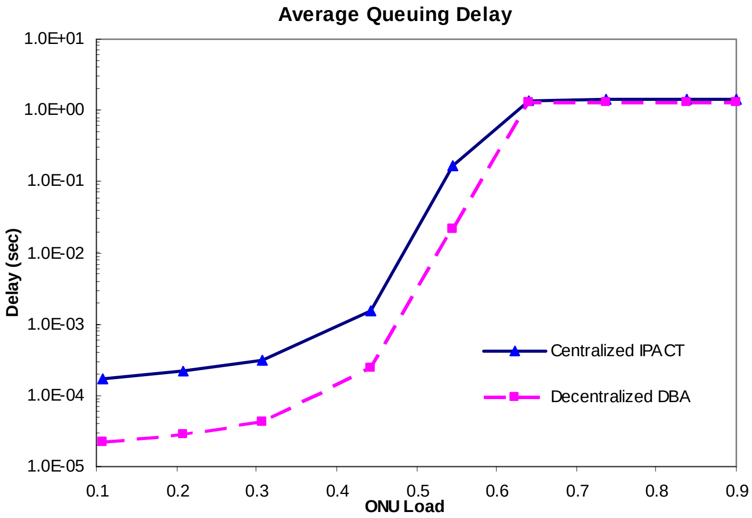

Fig. 2. Centralized versus decentralized: average queuing delay.

As the load increases beyond 0.5, the delay difference between the two schemes changes progressively from microseconds to milliseconds (queuing delay is now significant), mainly for the following reason: As the offered ONU load increases, more and more ONUs may request bandwidth greater than B_{MAX} (heavily loaded ONUs), while some other ONUs may still request less than B_{MAX} (lightly loaded ONUs). In this case, in contrast to the centralized limited scheme, the proposed DBA scheme reallocates unclaimed cycle bandwidth to heavily loaded ONUs [see Eq. (4)]. This results in more efficient upstream channel utilization and consequently minimizes the average packet queuing delay. Finally, as the network saturates $(\text{OOL} > 0.6)$, the performances of the two schemes are almost identical. This is because most ONUs now request more than the predetermined maximum allowed bandwidth (B_{MAX}), and thus most ONUs get the same allocation (B_{MAX}). This, in turn, eliminates the main advantage at higher loads, which is the reallocation of the unclaimed cycle bandwidth.

The results of Fig. 3 corroborate the results of Fig. 2, since lower average packet queuing delay indicates more efficient upstream channel utilization. Eliminating guard band slots increases the available upstream channel bandwidth, and redistributing unclaimed bandwidth of lightly loaded ONUs to highly loaded ONUs leads to more efficient channel utilization.

4.B. Performance of Upstream Traffic Including Both LAN and MAN/WAN

In this section, we divide upstream traffic into 20% LAN traffic and 80% MAN/WAN traffic. Figures 4, 5 compare the average and maximum end-to-end packet delay of LAN traffic as a function of OOL for both centralized IPACT and distributed DBA schemes. As

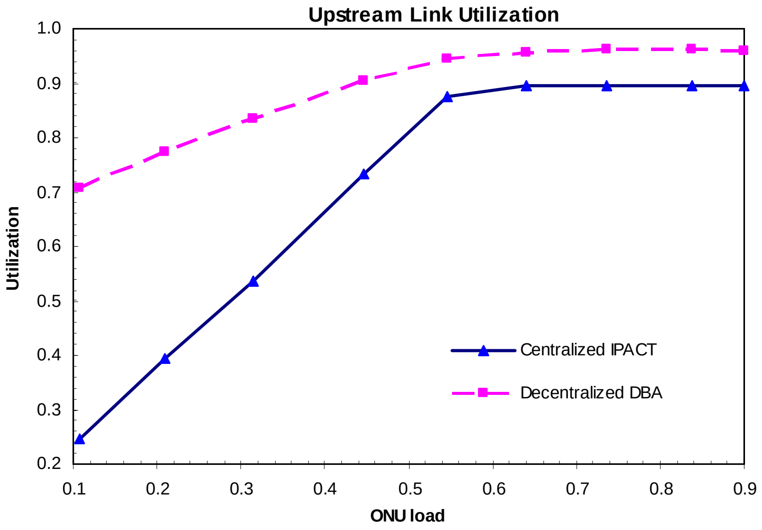

Fig. 3. Centralized versus decentralized: upstream link utilization.

can be seen from both figures, at all network loads the average and maximum end-to-end packet delays of the distributed scheme are always less than those of the centralized one. As can be seen from Fig. 4, at low load the centralized scheme has approximately 250–350 µs longer average end-to-end delays than that of the distributed one. This is due mainly to the round-trip-time (RTT) delay from the ONUs to the OLT (more than 210 µs for a 20 km trunk). In the case of the proposed ring-based architecture, the maximum propagation delay between the two most distant ONUs on the ring is approximately 15 µs. At higher load, as queuing delay becomes higher and dominant, the delays get higher, and difference between the two schemes are now in milliseconds.

5. Conclusions

This paper has proposed a novel ring-based local access PON architecture that has addressed some of the limitations of current tree-based PON architectures that include supporting private networking capability. Specifically, we have proposed and devised a simple ring-based EPON architecture that supports a truly shared LAN capability as well as upstream access to the OLT. The proposed architecture supports a fully distributed control plane among the ONUs for ONU–ONU communication as well as upstream access to the OLT. A distributed DBA scheme that supports the proposed decentralized architecture has been developed, and its performance has been compared with that of a tree-based centralized scheme. The simulation results have indicated that the overall performance of the proposed distributed scheme, including average packet queuing delay and channel utilization, outperforms that of the centralized one, particularly for LAN traffic.

While the proposed architecture slightly increases the complexity and the cost of the ONU, its main feature of supporting a truly shared LAN capability among end users within a PON-based local access infrastructure might justify the extra cost. In addition, the proposed architecture offers several key advantages over typical star-based centralized PON architectures: (1) Since the OLT always receives a fully regenerated upstream signal from the last ONU, upstream power levels received by the OLT are always guaranteed to be almost constant. This eliminates the typical near–far problem and, consequently, the need for a burst-mode receiver at the OLT. (2) Regeneration and retransmission of upstream traffic at

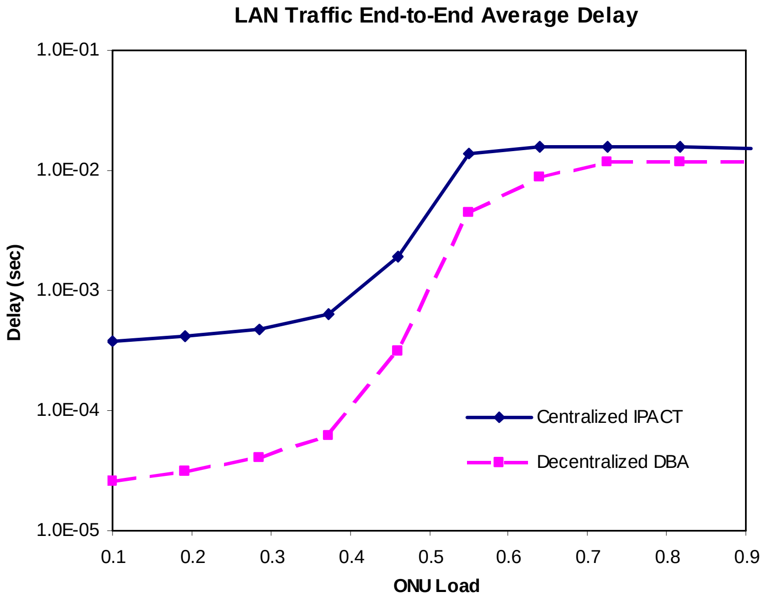

Fig. 4. Centralized versus decentralized: end-to-end average delay of LAN traffic.

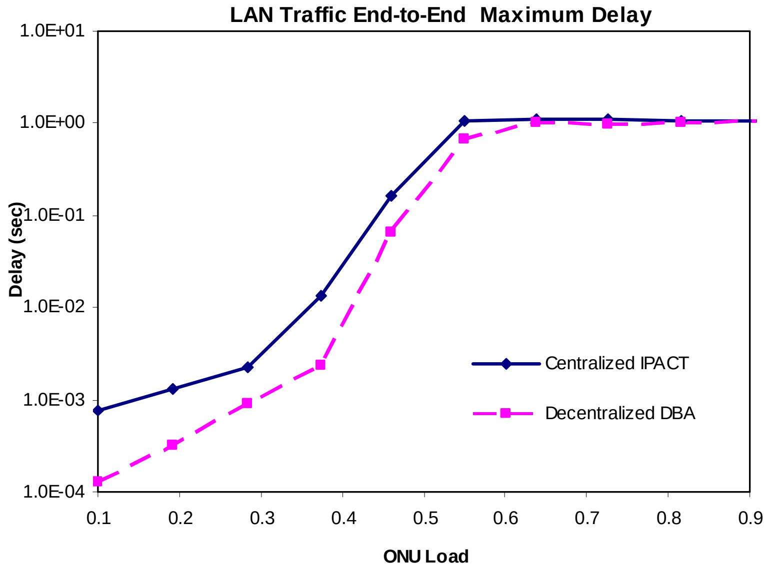

Fig. 5. Centralized versus decentralized: end-to-end maximum delay of LAN traffic.

each node ensures the following: (a) Collision-free upstream data transmission is obtained without resort to the typical use of guard time bands. This saving of guard time's overhead increases the available upstream transmission bandwidth and minimizes queuing delay. (b) The signal level attains full power at every node. This eliminates the typical limited-power-budget problem associated with the physical layer LAN emulation techniques described above.

References and Links

[1] B. Lung, "PON architecture 'future proofs' FTTH," Lightwave **16**(10), 104–107 (1999).

[2] Alloptic, "Ethernet Passive Optical Networks," The International Engineering Consortium, http://www.iec.org

[3] G. Kramer, B. Mukherjee, and G. Pesavento, "Ethernet PON: design and analysis of an optical access network," Photon. Network Commun. J. **3**(3) (2001).

[4] G. Kramer and G. Pesavento, "Ethernet Passive Optical Network (EPON): Building a Next Generation Optical Access Network," IEEE Com. Mag., pp. 66–73, February 2002.

[5] IEEE 802.3ah Task Force [Online]. Available: http://www.ieee802.org/3/efm/public/index.html

[6] D. Sala, B. Gaglianelo, A. Gummalla, Y. Goff, J. Limb, E. Nedellec, C. Ribeiro, and G. Rishi, "IEEE EFM," May 2002 materials, www.ieee802.org/3/efm.

[7] C.-J. Chae, S.-T. Lee, G.-Y. Kim, and H. Park, "A PON system suitable for internetworking optical network units using a fiber Bragg grating on the feeder fiber," IEEE Photon. Technol. Lett. **11**, 1686–1688 (1999).

[8] E. Wong and C.-J. Chae, "CSMA/CD-based EPON with optical internetworking capability among users," IEEE Photon. Technol. Lett. **16**, 2195–2197 (2004).

[9] G. Kramer, B. Mukherjee, and G. Pesavento, "Interleaved Polling with Adaptive Cycle Time (IPACT): a dynamic bandwidth distribution scheme in an optical access network," Photon. Network Commun. **4**, 89–107 (2002).

[10] C. Assi, Y. Ye, S. Dixit, and M. Ali, "Dynamic bandwidth allocation for quality of service over Ethernet PONs," IEEE J. Select. Areas Commun., December 2003.

[11] D. Nikolova, B. Van Houdt, and C. Blondia, "Dynamic bandwidth allocation algorithm in EPON: a simulation study," presented at Opticomm'03, Dallas, Texas, October 2003.

[12] A. Shami, X. Bai, C. Assi, and N. Ghani, "Jitter performance in Ethernet Passive Optical Networks," J. Lightwave Technol. **23**, 1745–1753 (2005).

[13] A. Shami, X. Bai, N. Ghani, C. Assi, and H. T. Mouftah, "QoS control schemes for two-stage Ethernet Passive Optical Access Networks," IEEE J. Select. Areas Commun. **28**, 1467–1478 (2005).

[14] V. Paxson and S. Floyd, "Wide area traffic: the failure of Poisson modeling," IEEE/ACM Trans. Netw. **3**, 226–244 (1995).

[15] W. Willinger, M. S. Taqqu, and A. Erramilli, "A bibliographical guide to self-similar traffic and performance modeling for modern high-speed networks," in *Stochastic Networks* (Oxford U. Press, 1996), pp. 339–366.