A Novel Ring-Based WDM-PON Access Architecture for the Efficient Utilization of Network Resources

H. Erkan, ASM D. Hossain, R. Dorsinville, and M.A.Ali
Dept. of Electrical Engineering
City University of New York, NY, USA

G. Ellinas
Dept. of Electrical and Computer Engineering
University of Cyprus, Nicosia, Cyprus

A. Hadjiantonis
Dept. of Electrical Engineering
Intercollege, Nicosia, Cyprus

A. Khalil
Engineering Science Dept.
LaGuardia College, CUNY, New York, NY, USA.

Abstract—This work proposes a simple and cost effective local access WDM-PON architecture that combines the salient features of both traditional static WDM-PON (i.e., dedicated connectivity to all subscribers with bit rate and protocol transparencies, guaranteed QoS, and increased security) and dynamic WDM-PON (i.e., efficiently utilizing network resources via dynamic wavelength allocation/sharing among end-users). Specifically, this paper proposes and devises a novel ring-based local access WDM-PON architecture that efficiently supports dynamic allocation of wavelengths/timeslots and sharing traffic as well as a truly shared LAN capability among PON end-users.

I. INTRODUCTION

Passive Optical Network (PON) technology is emerging as a viable solution for next-generation broadband access networks [1-5]. A PON connects a group of Optical Network Units (ONUs) located at the subscriber premises to an Optical Line Terminal (OLT) located at the service provider’s central office (CO). Among the various PON schemes, single channel Time-Division Multiplexed PON (TDM-PON) and multi-channel Wavelength-Division Multiplexed PON (WDM-PON) architectures are the two most viable candidates. TDM-PON supports a single wavelength channel in the downstream direction (OLT to ONUs) and another wavelength in the upstream direction (ONUs to OLT).

WDM-PONs are emerging as the most promising future access solutions that can provide evolutionary upgrade to existing TDM-PONs. These schemes can support multiple wavelengths in either or both the upstream and downstream directions. In traditional WDM-PON, each subscriber (ONU) is assigned a separate pair of dedicated upstream and downstream wavelength channels. In addition to its operational simplicity, this approach provides dedicated point-to-point optical connectivity to each subscriber with bit rate and protocol transparencies, guaranteed QoS, and increased security.

Despite these aforementioned numerous crucial advantages, typical static WDM-PON architectures suffer from several limitations including inability to efficiently utilize network resources and to cope with the dynamic and bursty traffic patterns of the emerging integrated triple play services [5-6]. The former limitation is exacerbated when some wavelength channels are heavily loaded while others are underutilized or are totally idle. In this case, unused dedicated channel capacities of those lightly loaded/idle subscribers cannot be shared by any of the other heavily loaded users attached to the PON, leading to the waste of scarce network resources. Therefore, to increase the total throughput, future WDM-PON architectures must support dynamic bandwidth allocation (DBA) and sharing.

To address these problems, several WDM-PON architectures and protocols that dynamically manage and allocate bandwidth in both time and wavelength dimensions have been proposed recently [6-10]. Most of these schemes, however, are costly and assume complex OLT and ONU setups which require tunable, or arrays of fixed transceivers or both, WDM filters, and wavelength-band-selective receivers. Furthermore, schemes that support dynamic wavelength sharing [6-7], where additional wavelength channels are added to accommodate the fraction of bursty downstream traffic that may exceed the user’s dedicated downstream wavelength channel rate, are still falling short of addressing the fundamental problem of the inefficient utilization of network resources. This is because the unused capacities of those lightly loaded/idle dedicated downstream wavelength channels are still being wasted.

For WDM-PONs to evolve as next generation access solutions, they must support all of the features offered by existing TDM-PONs, including simplicity, flexibility, cost-effectiveness, dynamic bandwidth allocation, and private networking capability. In this regard, the emulation of a shared Local Area Network (LAN) within a single PON has recently received considerable attention, but only within the context of TDM-PONs [11-13].

The purpose of this paper is to propose a simple and cost effective local access WDM-PON architecture that combines the salient features of both traditional static WDM-PON (i.e., dedicated connectivity to all subscribers with bit rate and protocol transparencies, guaranteed QoS, and increased security) and dynamic WDM-PON (i.e., efficiently utilizing network resources via dynamic wavelength allocation/sharing among end-users). Specifically, this paper proposes and
devises a novel ring-based local access WDM-PON architecture that efficiently supports dynamic allocation of wavelengths/timeslots and sharing traffic as well as a truly shared LAN capability among ONU end-users.

Unlike a typical WDM metro-access ring network, where the feeder fiber of a PON is replaced with a metro fiber ring that interconnects the hub and access nodes, the proposed architecture interconnects WDM ONUs via a short distribution fiber ring in the local loop but allows them to share the feeder fiber for long reach connectivity to the OLT. This architecture is well suited for an autonomous access environment such as a university campus or a private corporation where several high-end users are closely dispersed within a 1-4 km diameter area.

To the best of our knowledge, this is the first WDM-PON architecture that: 1) dynamically allocates unused capacities of lightly loaded/idle wavelength channels to heavily loaded channels in a simple WDM-PON configuration that typically assigns a dedicated wavelength channel to each subscriber; and 2) supports a truly shared LAN capability among WDM-PON end-users;

The rest of the paper is organized as follows: Section II describes the proposed architecture and Section III discusses the LAN DBA scheme under a fully distributed control plane. Section IV presents the dynamic hybrid time/wavelength-based scheduling algorithm and Section V presents simulation results.

II. PROPOSED ARCHITECTURE

Figure 1 illustrates the proposed ring-based WDM-PON architecture. An OLT is connected to N WDM ONUs (this work assumes N = 16) via a 20-km trunk feeder fiber, a passive 3-port optical circulator, and a small fiber ring. To cover the same local access area as in the tree-based architecture [1-5], the small ring at the end of the trunk is assumed to have a 1-4 km diameter. The ONUs are joined by point-to-point links in a closed loop around the access ring. The links are unidirectional: both downstream and upstream signals (combined signal) are transmitted in one direction only. Each ONU is assigned a single dedicated wavelength channel, $\lambda_{\text{LAN}}$, which is terminated, regenerated, and retransmitted at each ONU.

The OLT houses an array of N fixed transmitters (Tx) and another array of N+1 fixed receivers (Rx), a passive 3-port optical circulator, a flow scheduler, and a commercially available low cost thin film-based DWDM multiplexer/demultiplexer with channel dependent insertion losses between 0.8 and 3.8 dB. Each Tx/Rx pair corresponds to one ONU and utilizes the same wavelength for transmitting and receiving downstream and upstream traffic, respectively. The extra receiver (N+1) located at the OLT is used to detect the local control/LAN channel. Each ONU has a Tx/Rx pair which is matched to the corresponding pair at the OLT and another Tx/Rx pair for transmitting and receiving the local LAN channel, $\lambda_{\text{LAN}}$. In addition, each ONU houses a commercially available low cost four-port thin film filter-based fixed optical add-drop multiplexer (OADM), where two wavelengths (corresponding dedicated downstream/upstream and LAN wavelengths) are dropped and added at each node.

The DWDM downstream signal is coupled to the ring via port 3 of the optical circulator. After recombining it with the re-circulated LAN signal via a 2x1 WDM combiner (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/add and go-through fashion. For instance, at the first ONU, the dedicated downstream wavelength channel $\lambda_1$ along with the re-circulated control/LAN channel are dropped and processed; then, the dedicated upstream signal $\lambda_1$ along with the regenerated control/LAN channel are added. Finally, at the last node (ONU$_N$), wavelengths $\lambda_N$ and $\lambda_{\text{LAN}}$ are dropped/added. Thus, the DWDM downstream signal is terminated at the last node.

The combined DWDM upstream and LAN signal emerging from the last ONU at the end of the ring is split into two components via a (10:90) 1x2 passive splitter placed on the ring directly after the last ONU. The first component (90 percent) is directed towards the OLT via circulator ports 1 and 2, while the second component (10 percent) passes first through a band rejection filter that terminates the DWDM upstream signal. The LAN signal emerging from the band rejection filter is allowed to re-circulate around the ring after recombining with the downstream signal (originating from the OLT) via the 2x1 WDM combiner. The first component of the combined DWDM upstream and LAN signal is received and processed by the array of N+1 fixed optical receivers (housed at the OLT). Specifically, each of the N upstream optical receivers detects the corresponding upstream signal and recovers the MAN/WAN traffic, while the LAN optical receiver, as will be explained below, processes the control messages and may discard or process the LAN traffic.

III. LAN DBA SCHEME UNDER A FULLY DISTRIBUTED CONTROL PLANE

Direct intercommunication among ONUs is implemented via a LAN/control wavelength channel, which is terminated, regenerated, and retransmitted at each ONU. Both control messages (REPORTs) and LAN data share the same
wavelength channel (in-band signaling). Since REPORT messages along with LAN data are processed and retransmitted at each node, ONUs can directly communicate their LAN queue status and exchange signaling and control message information along with LAN traffic with one another. The REPORT message typically contains the desired size of the next timeslot based on the current ONU’s LAN buffer occupancy.

The proposed LAN distributed scheme utilizes a time division multiple access (TDMA) arbitration scheme in which the OLT is excluded from the arbitration process. The algorithm is cycle-based, where a cycle, $T_{LAN}$, 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 $T_{MAX}$. Each ONU maintains a database about the state of each other ONU’s LAN queue. This information is updated each cycle whenever the ONU receives new REPORT messages from all other ONUs. During each cycle, the ONUs sequentially transmit their REPORT messages along with LAN data in an ascending order within their granted LAN timeslots around the ring from one node to the next, where each REPORT message is finally removed by the source ONU after making one trip around the ring.

Note that in this work, as will be shown below, LAN data might be a combination of native local LAN traffic and transient downstream (TDS) traffic. TDS traffic is defined here as downstream traffic (the fraction of bursty downstream traffic that may exceed the user’s dedicated downstream wavelength channel rate) destined to a given access node ONU$_i$ but terminated at a different transient access node ONU$_j$. This is because TDS traffic is always transported via a non-dedicated lightly loaded downstream wavelength channel $\lambda_j$ other than its own dedicated heavily loaded downstream wavelength channel $\lambda_i$, which is typically pre-assigned to transport native ONU$_i$’s downstream traffic. Once the TDS traffic is terminated at ONU$_j$, it is handled as transient LAN traffic whose new source is ONU$_i$ and final destination is ONU$_j$. The TDS traffic along with ONU$_j$’s local LAN traffic are first buffered at the corresponding ONU$_j$’s LAN queue, and are then retransmitted over the LAN/control wavelength channel around the ring to their final destinations, within the same granted ONU$_j$’s LAN timeslot.

The distributed LAN DBA algorithm utilized here is the same as that reported in [14] for a single channel TDM-PON, but modified here to account for the TDS traffic. The reader is referred to [14] for further details regarding the algorithm. The algorithm provisions native LAN data as well as TDS traffic, but handles these two different forms of data as LAN data and doesn’t distinguish between them. Thus, when an ONU reports its LAN queue status to other ONUs as well as to OLT, it must periodically report the aggregate bandwidth $(Q_{Report, Aggr})$ of all data traffic buffered at its LAN queue without any distinction, including its own local LAN traffic as well as TDS traffic (if there is any). The DBA module housed at each ONU uses this information in each cycle to calculate one new LAN timeslot assignment for each ONU. Since LAN transmission is based on a TDMA scheme, inter-ONU traffic including local LAN data and control messages as well as TDS traffic, are all transmitted within the same granted transmission window.

IV. DOWNSTREAM WAVELENGTH ASSIGNMENT & SHARING

The OLT houses 16 downstream queues, each queue, $Q_i^{OLT}$, is assigned to a specific ONU$_i$ and is connected to a dedicated downstream wavelength, $\lambda_i$. Each ONU houses two queues, one queue, $Q_{i,up}^{ONU}$, is assigned to a dedicated upstream wavelength, $\lambda_i$, while the other queue, $Q_{i,LAN}^{ONU}$, is assigned to the LAN/control traffic. The process of dynamically assigning/sharing downstream wavelengths is implemented jointly at both the OLT and the ONUs as follows: If a dedicated downstream wavelength channel, $\lambda_i$, with traffic
destined to ONUj is overloaded (i.e., incoming bursty traffic flows may exceed the dedicated channel rate for some interval, so that its corresponding $Q^{OLT}_{i}$ is congested), the following steps are executed:

1. The scheduler at the OLT searches for another underutilized/idle downstream wavelength channel, $\lambda_{j}$, i.e., a channel whose corresponding queue, $Q^{OLT}_{j}$, has some available space that can accommodate one or more of $\lambda_{j}$’s excess flows. This queue will be referred to as an available queue.

2. If the search is successful, the available channel, $\lambda_{x}$, is selected if and only if its corresponding LAN queue at ONUj, $Q^{ONU}_{j,LAN}$, is also available.

3. The scheduler redirects one, some, or all of $\lambda_{x}$’s excess flow(s) arriving from the MAN/WAN to $Q^{OLT}_{j}$, where it is then transmitted, along with ONUj’s native downstream traffic arriving from WAM/MAN, to ONUj over its dedicated wavelength channel $\lambda_{j}$. Such a channel that can accommodate and transport, in addition to its own local downstream traffic, other wavelength channel’s downstream traffic, will be referred here as an “Acceptor Wavelength Channel (AWC)” and its corresponding queue as an “available ready queue”.

4. The $\lambda_{j}$-downstream optical receiver housed at ONUj terminates all $\lambda_{j}$’s downstream traffic including both native downstream traffic destined to ONUj and TDS traffic destined to ONUi, examines the destination MAC address of each detected Ethernet frame, and then performs the following two functions: i) native downstream traffic that matches ONUj’s MAC address is copied and delivered to the end-users; ii) TDS traffic destined to ONUi (whose MAC address does not match that of ONUj) is redirected to ONUj’s LAN queue. It is then transmitted, along with ONUj’s own local LAN traffic, as LAN traffic around the ring over $\lambda_{LAN}$ to their final destinations, within the proper designated LAN timeslot of ONUj.

5. TDS flows are returned back to their original queues once these queues are available.

Downstream queues located at the OLT are classified into three different types: a donor queue, an acceptor queue, and a regular queue. A donor queue is defined here as a queue that had some of its own native flows rerouted to other dedicated downstream queue(s). An acceptor queue is a queue that is buffering in addition to its own native local flows, one or more TDS flows that originally belong to other different congested queue(s). A regular queue is neither a donor nor an acceptor queue.

**IV.1 Shared Wavelength Selection & Scheduling (SWS) Algorithm**

In this section, a resource allocation scheme that efficiently supports dynamic allocation of wavelengths and sharing of traffic among PON end-users is developed. Specifically, we develop a shared wavelength selection and scheduling (SWS) algorithm that dynamically, as well as fairly and efficiently allocates downstream wavelength channels among end-users. To achieve this objective, downstream network resources (downstream wavelength channels) are load-balanced and efficiently utilized via traffic-engineered routing of end-user’s downstream traffic flows. Although the proposed scheme utilizes an OLT-based centralized control plane, its implementation requires the participation of the distributed LAN control plane.

The SWS algorithm is divided into two overlapping processes: The first process, shown in Figure 2, is a fixed periodic polling-cycle operation, where the flow scheduler at the OLT periodically checks the status of each dedicated downstream $Q^{OLT}_{i}$ and its corresponding $Q^{ONU}_{i,LAN}$. The second is the process of admitting/routing/dropping newly arriving native flow(s), which is triggered by the arrival of new native flow(s). The algorithm module maintains two databases. The first database keeps track of the records of all existing and newly arriving flows including flow ID, flow average rate, flow maximum rate, native flow’s pre-assigned dedicated wavelength channel, and TDS flow’s AWC. The second database maintains a Transient Counter (TCi) for each dedicated downstream queue.

The function of a TC is to keep track of the number of native flows that have been redirected from a given donor queue to other available queue(s) as well as the number of non-native TDS flows that are being buffered by an acceptor queue. The number displayed by a TC can either be a zero, positive integer (to indicate that this queue has had a surplus capacity), or a negative integer (to indicate that this queue has had a capacity deficit), corresponding to a regular queue, an acceptor queue, and a donor queue, respectively. For instance, if TCi = -2, this indicate that $Q^{OLT}_{i}$ is a donor queue that had two of its native flows routed to other queue(s). These databases are updated periodically in each polling cycle upon the arrival of new flows or the termination of an existing flow.

**A. First Process: A Periodic Polling-Cycle Operation**

This operation is triggered at the end of each polling cycle, where the cycle time is assumed to be fixed ($T_{polling}$). The criterion used to determine whether a downstream queue is congested or available is based on the available size (capacity) of that queue at the instant of polling cycle scheduling. This is equivalent to utilizing a dynamic threshold versus typical static ones used in [6-7]. In this work, maximum available queue size is set such that the worst case packet delay at a given queue does not exceed 10 ms. Thus, for a 1Gb/s downstream output queue rate (all network DWDM channel capacities are assumed to be equal including downstream, upstream, and LAN channels, each operating at a 1 Gb/s), the maximum available queue size is (1Gb/s) (10 ms) = 10 Mbits. If the available size of $Q^{OLT}_{i}$ at the end of the polling cycle cannot accommodate the accumulated average-sum rate of all existing downstream...
flows during the next polling cycle n+1, \( Q_{OLT}^{i} \) is congested. In other words, if the following inequality does not hold, \( Q_{OLT}^{i} \) is congested:

\[
\sum_{j} f_{j, \text{peak}}^{\text{exist}} - R_{out}^{\text{OLT}} \leq \sum_{k} f_{k, \text{peak}}^{\text{new}} - R_{out}^{\text{OLT}} \leq Q_{OLT}^{i, \text{ava}}(t)
\]

(1)

where \( f_{j, \text{peak}}^{\text{exist}} \) is the average rate of an existing flow j, \( R_{out}^{\text{OLT}} \) is the downstream output queue rate, \( T_{p} \) is the polling cycle time, and \( Q_{OLT}^{i, \text{ava}} \) is the available size of queue i at the instant of polling cycle scheduling. If the inequality of Eq. 1 holds, this is just an indication that \( Q_{OLT}^{i} \) is not congested but does not necessarily guarantee that it is an available queue. For \( Q_{OLT}^{i} \) to be an available queue, its available size must accommodate the worst case peak-sum rate of all existing flows as well as, at least, one newly arriving flow at its peak rate. Thus, if the following inequality holds, \( Q_{OLT}^{i} \) is an available queue:

\[
\sum_{j} f_{j, \text{peak}}^{\text{exist}} + \sum_{k} f_{k, \text{peak}}^{\text{new}} - R_{out}^{\text{OLT}} \leq Q_{OLT}^{i, \text{ava}}(t)
\]

(2)

where \( f_{j, \text{peak}}^{\text{exist}} \) is the peak rate of an existing flow j and \( f_{k, \text{peak}}^{\text{new}} \) is the peak rate of a newly arriving flow K or a rerouted TDS flow. To avoid frequent flow transitions that might degrade the performance [6], the average rate of an existing flow in Eq. 1 has been replaced by the peak rate in Eq. 2, so that the difference between the available capacity of a non-congested queue and that of an available one is significant.

Similarly, the inequality used to determine the availability of a LAN queue is analogous to that used above in Eq. 2 for a downstream queue but appropriately modified to account for all the different input data contents of the LAN queue, including its own local LAN traffic as well as the TDS traffic (if there is any). If the following inequality holds, \( Q_{\text{LAN}}^{i} \) is an available queue:

\[
R_{\text{in,MAX}} + \sum_{j} f_{j, \text{peak}}^{\text{exist}} T_{p} - R_{\text{out,LAN}} \leq Q_{\text{LAN}}^{i, \text{ava}}(t)
\]

(3)

where \( R_{\text{in,MAX}} \) is the maximum access LAN link data rate from users to ONU, \( f_{j, \text{peak}}^{\text{exist}} \) is the peak rate of an existing TDS flow j that is utilizing ONU, \( f_{k, \text{peak}}^{\text{newTDS}} \) is the peak rate of a newly arriving TDS flow K, \( R_{\text{out,LAN}} \) is the output rate of the LAN queue, and \( Q_{\text{LAN}}^{i, \text{ava}} \) is the estimated available size of the LAN queue calculated at the instant of polling cycle as flows;

\[
Q_{\text{LAN}}^{i, \text{ava}} = Q_{\text{LAN,MAX}} - Q_{\text{Report,Aggr}}
\]

(4)

It is important to emphasize that in contrast to the actual accurate available size of a downstream queue that is used in Eq. 2 above, the available size of the LAN queue used in Eq. 3 is roughly calculated by the SWS algorithm module using Eq. 4. The SWS algorithm module located at the OLT periodically (in each LAN cycle) checks the status of each distant ONU’s LAN queue (\( Q_{\text{Report,Aggr}} \)) via the LAN control messages transmitted to the OLT. Thus, the state of the LAN queues used by the SWS algorithm to determine the available size of each distant LAN queue may not reflect the actual instantaneous status of the LAN queues.

At each polling cycle, the algorithm uses the most updated ONU’s LAN REPORT message, i.e., the REPORT message that arrives last to the OLT within the current polling cycle. Thus, Eq. 3 is an approximation and is not as accurate as Eq. 2. However, since Eq. 3 assumes the worst case scenario for all existing and incoming traffic, it is a sound approximation. This process requires that the fixed polling cycle period, \( T_{\text{polling}} \) of the SWS algorithm, be at a minimum equal to or greater than the maximum varying length of the DBA algorithm’s LAN cycle, \( T_{\text{LAN}} \).

As can be seen from the right part of Figure 2, once \( Q_{OLT}^{i} \) has been identified as a congested queue, the status of its corresponding TC is checked to determine its type. If \( Q_{OLT}^{i} \) is either a donor or a regular queue (\( TC_{i} \leq 0 \)), the algorithm determines how many flows should be removed from the queue (rerouted and/or dropped) so that the inequality of Eq. 1 holds, i.e., \( Q_{OLT}^{i} \) is no longer congested. Note that the actual process of rerouting these excess flows from the congested donor/regular queue to an available ready queue and/or dropping them is postponed until all other remaining queues in the polling cycle are also inspected. Alternatively, if \( Q_{OLT}^{i} \) is an acceptor queue (\( TC_{i} > 0 \)), i.e., a queue that is buffering one or more non-local TDS flows, the algorithm first attempts to redirect one or more of these excess TDS flows back to their original source queues (provided that these source queues are available) in a LIFO order until the inequality of Eq. 1 holds. Otherwise, these TDS flows are dropped. Note that in this case, however, an original source queue that restores its own native flows must be an available queue only (i.e., a queue that satisfies only the inequality of Eq. 2) and need not be an available ready queue, i.e., a queue that satisfies inequalities of both Eqs. 2 and 3.

The left part of Figure 2 shows the steps involved at the end of a given polling cycle n to determine whether a non-congested queue can also be qualified as an “available ready queue”; i.e., a queue, which is ready to buffer, during any instant within the next polling cycle n+1, one or more non-local TDS flows that belong to other queue(s). Once \( Q_{OLT}^{i} \) has been identified as a non congested queue (satisfies inequality of Eq. 1), its status is further examined to determine if it qualifies as an available queue and if it does,
the status of its corresponding TC$_i$ is checked to determine its type. If it is either a regular or an acceptor queue (TC$_i \geq 0$), the status of its corresponding LAN queue, Q$^{ONU}_{LAN}$, is examined to determine whether it is an available queue as well; if Q$^{ONU}_{LAN}$ is an available queue (satisfies inequality of Eq. 3), Q$^{i,OLT}$ is labeled as an “available ready queue”. Otherwise, if Q$^{i,OLT}$ is an available donor queue (TC<0), the algorithm first attempts to restore one or more of its native flows (which have been buffered at other downstream queues), until the queue is no longer available (inequality of Eq. 2 no longer holds). Thus, if a queue is an available donor queue, the process of restoring its own native flows has precedence over that of accepting and buffering non local TDS flows that belong to other queues.

B. Second Process: Admission & Rerouting of Newly Arriving Native Flows

When a new flow arrives at the OLT from MAN/WAN to the corresponding dedicated downstream queue, the status of this queue is examined to determine whether it is an available queue or not. If it is (satisfies inequality of Eq. 2), the flow is admitted and resources are allocated. If Eq. 2 is not satisfied, then the flow scheduler initiates a quick search to find an available and ready queue. If it finds one, it redirects the newly arriving flow to that queue, otherwise the newly arriving flow will be dropped.

C. Traffic-Engineered Selection Criteria

During each polling cycle, as can be seen from Figure 2, the algorithm may identify one or more regular/donor congested queues as well as one or more available ready queues. Depending on available network resources, a scenario may arise in which not all of the identified congested queues are able to reroute their excess flows to other available ready queues. Another scenario is that there might be only one congested queue but several other available ready queues. Thus, the algorithm, based on some selection criteria, must select which congested queue(s) from amongst all of the identified ones gets to reroute its excess flows; and which available ready queue(s) is to be selected from amongst several available ready queues. These criteria must take into account not only maximizing the network throughput, but also being fair to end-users; that is, these criteria must balance the trade-off between maximizing throughput and fairness. The selection criteria are:

a) To achieve fairness among end-users, the queue with the least number of exported native flows is selected first. Thus, a regular queue with a zero TC is the queue that is selected to route its excess native flows to other available ready queue(s) first; followed by the donor queue whose TC displays the minimum negative number, etc. Thus, the donor queue whose TC displays the maximum negative number gets to route its excess native flows last.

b) To maximize downstream throughput, the algorithm selects the least-loaded pair of available downstream and corresponding LAN queues at the OLT and at the ONU, respectively, i.e., the pair with the maximum available average capacity. In both cases, if there is a tie, one is randomly selected.

V. PERFORMANCE EVALUATION

An event driven WDM-PON simulator was developed using C++. The simulator consists of two modules. The first module executes a WDM/TDM-based downstream SWS algorithm at the OLT while the second module concurrently executes a TDM-based distributed LAN DBA at the ONUs. The traffic model used here is the same as that reported in [6-7], where bursty downstream flows are generated by
aggregating multiple sub-streams; each is modeled as an on/off source, with Pareto distribution to capture the self-similar nature of Ethernet traffic. Because the execution of the simulation model is computationally intensive and complex, a system with only 8 ONUs is used.

Based on the average down-link load per ONU, ONUs are classified into two different sets, with 4 ONUs each. In the first set, an average down-link load of 0.5 is destined to each ONU. In the second set, a variable down-link load that is incremented from 0.6 to 1.3 in 0.1 steps is destined to each ONU. Each ONU in the second set will be referred to as “ONU under test” (OUT). Average LAN traffic load per ONU is varied from 0.1 to 1.3 in 0.1 steps.

The system parameters assumed in the simulation are as follows: 1) line rates of all channels including downstream, upstream, and LAN channels, are all equal, each operating at a 1 Gbps; 2) each ONU houses a LAN buffer with 10 Mbits size; 3) the OLT houses 8 downstream buffers, each corresponding to a given ONU with 10 Mbits size; 4) downstream flow packets are standard variable Ethernet frames and arrive at the OLT from MAN/WAN, at a link rate of 2 Gbps for each dedicated downstream wavelength; 5) maximum access LAN link data rate from users to an ONU is 200 Mb/s; 6) maximum LAN cycle, $T_{LAN} = 2$ ms; 7) fixed polling cycle, $T_{polling} = 2$ ms; 8) downstream queues of the first set of ONUs has a fixed input of 10 flows each, while each downstream queue of the second set has a fixed input of 20 flows. For both sets, peak rate of each flow is set to 100 Mb/s; 9) the distance between the OLT and ONUs varies from 20 km to 23 km (ring circumference $\approx 3$ km); 10) all downstream traffic is treated as best effort traffic, i.e., all arriving frames at the OLT are queued at the corresponding downstream dedicated queues in a first-in-first-out (FIFO) order; 11) strict priority scheduling mechanism is assumed at each LAN queue; that is, LAN traffic is assumed to have higher priority compared to TDS traffic.

Figure 3 shows the down-link throughput per OUT as a function of both native LAN and OUT’s down-link loads. As can be seen from Figure 3, at very light LAN load and heavy down-link load, OUT’s down-link throughput is about 1.23 Gbps. Note that the theoretical OUT’s down-link throughput is given by: $\{\text{Dedicated downstream channel Rate (1 Gbps)} + \text{Unused LAN Capacity}\}/\text{Number of OUTs}$ (in this case number of OUTs is 4). Thus, under zero/very light native LAN load, the OUT’s maximum down-link throughput = 1.25 Gbps, which is in good agreement with the simulation results of Figure 3. As the LAN load increases, down-link throughput decreases. Finally, as expected, as the offered LAN load approaches unity or higher, down-link throughput levels-off at around 1 Gbps (dedicated channel rate). This is because transmission of TDS traffic is no longer possible.

**REFERENCES**