A Converged Optical Wireless Architecture for Mobile Backhaul Networks

Kostas Ramantas[1], Kyriakos Vlachos[1], Georgios Ellinas[2], and Antonis Hadjiantonis[3]
[1]Computer Engineering and Informatics Dept & Computer Technology Inst, Univ. of Patras, Greece
[2]Department of Electrical and Computer Engineering, University of Cyprus, Cyprus
[3]Department of Engineering, University of Nicosia, Cyprus
Email: {ramantas, kvlachos}@ceid.upatras.gr, gellinas@ucy.ac.cy, hadjiantonis.a@unic.ac.cy

Abstract—In this work we propose a new unified PON-RAN architecture for LTE mobile backhaul networks, employing ring-based WDM PONs. The proposed architecture supports dynamic setup of virtual circuits for inter-base station communication, over a dedicated λ_{LAN} channel. The reservation mechanism is arbitrated by the OLT, which also monitors the traffic imbalances of downstream channels. The proposed architecture also supports load balancing, by dynamically reallocating and sharing the capacity of the downstream wavelengths.

***Index Terms*— FiWi, Optical-Wireless Convergence, Optical Burst Switching, Passive Optical Networks (PONs)**

I. INTRODUCTION

Passive Optical Network (PON) technology has been revisited as a viable solution for next-generation fiber access networks [1]. The demand for high-access bandwidth is expected to grow continuously, due to the emergence of new real-time interactive applications with increased bandwidth requirements. This will inevitably require increasing manifold the access bandwidths offered by current generation PONs. Another challenge faced by operators is the need of backhauling traffic from next generation wireless broadband access architectures. The mobile backhaul, or radio access network (RAN), is utilized to transport traffic from individual base stations (BSs) to the access gateway (AGW). Legacy technologies such as circuit switched T1/E1 wireline or microwave used for existing 3G network infrastructures are not going to meet the capacity requirements of new 4G access architectures [2]. Thus, mobile operators are investing heavily in upgrading backhaul networks, migrating their infrastructure to fiber optic deployments ("Fiber To The Cell"). 4G base stations are expected to be densely populated to achieve high spectral efficiency and require high bandwidth and cost-effective backhauling. This makes Next Generation PONs (NG-PONs) a strong candidate for implementing the mobile backhaul due to the high capacity, longer range and economical deployment that they offer.

Many PON architectures have been studied in the literature. The most popular variant is TDM-based EPONs. In TDM-PONs a single wavelength is shared by all Optical Network Units (ONUs) in the upstream direction based on a TDMA algorithm which is arbitrated by the Optical Line Terminal (OLT). In order to increase access capacities, 10GE-PONs have been proposed and recently standardized [1]. They offer compatibility with pre-existing PON deployments and share existing passive components, resulting in a smooth network evolution. However, in new deployments, more efficient NG-PON architectures can be considered [3], that employ WDM technology to offer capacities beyond 10 Gbps and long reach albeit at a higher cost. WDM-PONs can be considered as an evolutionary scenario of existing TDM-PONs employing a dedicated wavelength per ONU for OLT/ONU communication.

Regarding next generation Wireless Broadband Access Networks (NG-WBANs), Mobile WiMAX and LTE are two competing technologies that are expected to achieve data rates beyond 100 Mb/s per end user. Specifically, LTE Advanced networks include the possibility for peak data rates up to 1 Gbit/s and latency less than 10ms [4]. Additionally, unlike WiMAX, LTE uses an evolution of the existing Universal Mobile Telecommunication System (UMTS) infrastructure, used by over 80% of the mobile operators [2]. Thus, it is not necessary to build a new network infrastructure, making LTE more popular with operators worldwide. Many works in the literature have focused on the integration of NG-PONs and NG-WBANs, to build a converged architecture that combines the merits of both technologies. Such converged infrastructure would enable the deployment of new, innovative high-bandwidth services, that support mobility and end-to-end QoS guarantees. One of the challenges that must be met by future converged architectures is the efficient inter-communication of LTE Base Stations. In [5], the authors propose a WDM-PON ring-based converged architecture which supports ONU inter-communication with a dedicated λ_{LAN} channel. In [6] a tree-based converged architecture is proposed, with the ability of all-optical communication of all BSs that belong to the same PON.

In this work we propose a new unified PON-RAN architecture for LTE mobile backhaul networks, employing ring-based WDM PONs. Mobile backhaul networks are a perfect candidate for exploiting the high capacity, inherent resilience, and ubiquity offered by WDM rings, as the – comparatively higher – infrastructure cost due to the use of WDM components is amortized to a much larger number of mobile clients. However, employing current generation

WDM-PON networks results in the inefficient use of resources, as wavelength capacity cannot be reallocated or shared between Optical Network Units (ONUs). Furthermore, the communication among ONUs is performed via the Optical Line Terminal (OLT), unnecessarily increasing delay and wasting capacity of both the upstream and the downstream channel. To solve these problems, the proposed architecture supports all-optical inter-communication and fully meshing of LTE Base Stations.

The rest of the paper is organized as follows. The new converged architecture is presented in Section II. In Section III a resource reservation protocol is detailed for setting up all-optical virtual circuits at the λ_{LAN}. In addition, a downstream wavelength sharing scheme is proposed, to support load balancing between ONUs/eNBs. Finally, in Section IV the proposed architecture is evaluated with extensive simulation experiments utilizing the ns-2 simulator framework.

II. Ring Architecture & ONU Design

The LTE network architecture consists of an all-IP core network, called the EPC (Enhanced Packet Core) and new, enhanced Base Stations called "evolved nodeBs (eNBs)". The eNBs are connected by means of the S1 interface to the EPC, whose logical components are the Mobility Management Entity (MME), the Serving Gateway (S-GW) and the Packet Data Network gateway (P-GW), together also known as the access gateway (AGW). LTE also introduced support for inter-BS connectivity via the X2 interface, to support handover operations. Recent studies have estimated traffic traversing the X2 interface to reach 4-10% of traffic traversing the S1 interface [6]. Thus, it is important for efficient converged architectures to support at least partial meshing of eNBs, so that X2 traffic does not flow through the AGW, which would waste resources and significantly increase packet delay. Furthermore, it is generally accepted that fiber deployment to cell towers ("Fiber-To-The-Cell") is the only future-proof solution to build mobile backhauls, which will scale to the increased capacity requirements of future NG-WBAN technologies [2]. Converged architectures based on PONs have the added benefit of reusing passive components after upgrading the active infrastructure, providing a cost-effective network upgrade path.

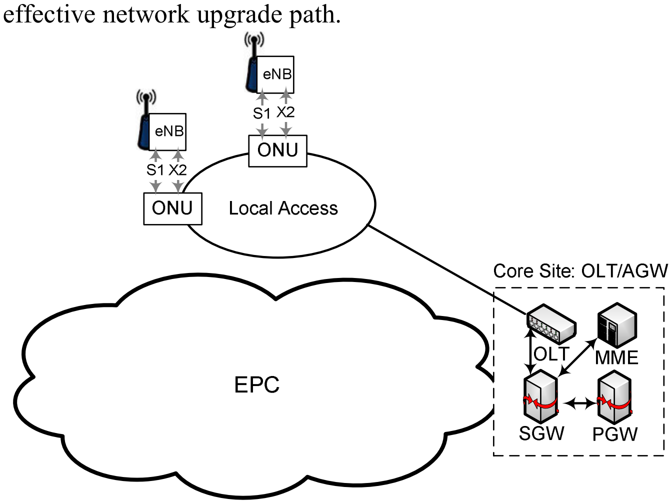

Figure 1: Proposed converged architecture

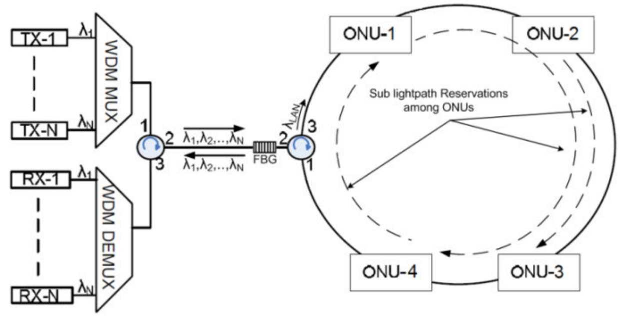

Figure 2: WDM-PON ring-based architecture

A. *Converged Architecture*

The proposed converged access architecture (Fig. 1) employs WDM rings of ONUs to interconnect LTE base stations (BSs) in the same local access area. Each WDM ring is connected to the corresponding Optical Line Terminal (OLT) via a bidirectional 10-20 km feeder fiber and a passive 3-port circulator. OLTs are collocated with the access gateways at the core sites of the Enhanced Packet Core. Each WDM ONU serves one LTE eNB, while both ONUs and eNBs are assumed to support a common standard interface. For example the eNB implementation may use two GigE interfaces to transport X2 and S1 interface traffic, or a single 10 GigE interface where X2/S1 traffic is multiplexed. These are interconnected with the corresponding Ethernet interfaces of WDM ONUs, which are collocated with the eNBs. Thus, traffic from the eNB is terminated at the WDM ONU and re-transmitted at the corresponding wavelength. Each ONU is assigned two wavelengths, namely a dedicated wavelength for downstream/upstream traffic from/to the OLT and another one, denoted as λ_{LAN} that is shared by all ONUs across the ring, for inter-ONU communication (see Fig. 2).

Unlike the converged architecture proposed in [5], a new ONU design allows bypassing intermediate ONUs, thus avoiding unnecessary termination of λ_{LAN} traffic in intermediate ONUs. The new ONU design supports all-optical meshing of eNBs that belong to the same ring, offering all-optical sub-wavelength connectivity [7]. Transmission is unidirectional at the ring: both upstream and downstream signals are transmitted in the same direction. A Fiber Bragg Grating (FBG) reflects back the λ_{LAN} wavelength from the upstream signal heading to the OLT and allows it to re-circulate around the ring.

Converged LTE architectures support end-to-end QoS by setting up logical links between the User Equipment (UE) and the Gateway, termed "EPS bearers". Each bearer corresponds to a unique QoS identifier (QCI) and specific QoS parameters (i.e., delay, loss, and bandwidth). Bearers are divided in two broad categories: GBR (Guaranteed Bit Rate) where blocking is preferred over packet dropping and Non-GBR where packet dropping may be experienced. Typically GBR traffic as well as non-GBR traffic that corresponds to carrier-provided services is distinguished and prioritized, so that it has the required packet forwarding treatment end-to-end. Since the EPS bearer is not visible at the backhaul network (i.e., ONUs and OLTs), the Gateway appropriately marks Layer 2 frames with a 3-bit Class-of-Service (CoS) value, carried at the "User Priority bits" of the Ethernet frame header. Lower priority non-GBR traffic (such as web browsing, P2P traffic, online video, etc) may suffer congestion-related losses at the backhaul network. In Section III we detail a load balancing scheme, which allows excess traffic from congested downstream channels to reach its destination ONU through one or more uncongested ones.

B. ONU Design

Figure 3 displays the ONU block design. Each ONU is equipped with a pair of lasers and receivers; λ_i is used for OLT-to-ONU downstream/upstream communication and λ_{LAN} for inter-ONU communication (termed as LAN traffic). Sub-wavelength sharing of the λ_{LAN} involves the aggregation of packets per destination ONU in a separate Virtual Output Queue (VOQ). Aggregated packets are then transmitted in burst mode over the λ_{LAN} wavelength. An 1x2 optical switch is used to either extract bursts at the destination ONU or transparently forward them to the next ONUs. Burst-mode transmission technology is very mature as it is already being used in commercial TDM PONs, where ONUs transmit bursts of packets during time slots pre-allocated by the OLT. By extracting bursts that reach their destination via the optical switch (destination stripping) the common wavelength can be spatially reused by downstream nodes, leading to an increased capacity. It must be noted that fast switches and burst-mode transceivers are also the building blocks of Optical Burst Switching (OBS), which has been extensively researched for sub-wavelength capacity provisioning in future backbone networks and metropolitan rings [8].

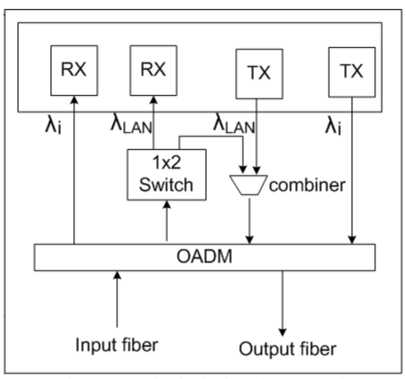

Figure 3: Block design for ONU-i

III. Dynamic Bandwidth Allocation / Reallocation

In the proposed architecture, inter-ONU communication is performed at the common wavelength λ_{LAN}. This allows full meshing of the LTE base stations that belong to the same ring, which can directly exchange traffic via the X2 interface. Additionally, this common wavelength can be used for offloading excess traffic from congested downstream channels, and re-directing it through one (or more) uncongested one (s). Thus, the network can react to short-term traffic changes by dynamically reallocating and sharing the capacity of the downstream wavelengths. Dynamic Bandwidth Allocation for LAN traffic is implemented at the OLT, which is aware of the utilization profiles of all wavelength channels.

A. LAN traffic reservation Protocol

PONs are typically associated with centralized control. For example, in TDM PONs ONUs are assumed to be synchronized with the OLT, which controls data transmission in the upstream direction using GATE and REPORT messages. Thus, in the proposed architecture, sub-wavelength connectivity at the λ_{LAN} is also arbitrated by the OLT. Inter-ONU communication is achieved by setting up all-optical virtual circuits on-demand between ONUs. For example, virtual circuits are setup between ONUs that serve neighboring Base Stations, which need to exchange handover traffic. In what follows, we detail the reservation mechanism used by the OLT to collect bandwidth requests from ONUs, schedule/setup virtual circuits over the λ_{LAN} channel and keep track of the reserved link capacity over the PON ring.

In the proposed reservation protocol, connection requests arrive at the OLT in the form of 3-tuples (s,d,c), where s is the originating ONU, d the destination ONU and c the bandwidth requirement. The OLT executes a scheduling algorithm to satisfy the request, by reserving capacity on every link across the path from node s to node d. It must be noted that the bandwidth requirement of handover traffic traversing the X2 interface (e.g., mean rate, peak rate, or effective bandwidth) can be estimated from traffic measurements performed by the corresponding ONU. A vital requirement of the converged architecture is the allocated capacity to satisfy quality-of-service (QoS) constraints, most notably transmission delay. For example, in LTE Advanced networks latency is expected to be less than 10 ms.

The proposed scheme utilizes Time Division Multiple Access (TDMA) arbitration with a fixed cycle time (t_{cycle}) to reserve capacity on the ring. Timeslots are allocated to the ONUs by the OLT, which implements a simple First-Fit scheduling algorithm. Time is discretized in fixed time-slots of duration τ, thus the utilization profile of each link is divided in a reservation window of K time-slots. The scheduling algorithm searches for a set of free consecutive slots on every link across the path from node s to node d that satisfy the connection request. If the reservation succeeds, the corresponding link capacity vectors are updated and ONUs are informed of the new reservation with appropriate signaling messages. Then, they synchronize the bypass operation of their 1x2 optical switches according to the designated timeslots, and schedule the transmission of packets in VOQs. In case that the reservation fails, the connection request will be blocked.

B. Downstream Capacity Sharing

In LTE networks load balancing is supported among Base Stations that are logically connected via the X2 interface, so that they can exchange traffic overloads. Thus, sub-wavelength λ_{LAN} connections can be setup from congested ONUs/eNBs to one or more uncongested ones, to facilitate downstream capacity sharing. For example, let us assume that λ_i is a heavily-loaded wavelength channel dedicated to ONU_i and λ_k a lightly-loaded channel dedicated to ONU_k. A fraction of the downstream traffic destined to ONU_i/eNB_i can be transported via λ_k and terminated at ONU_k. This traffic sub-flow, termed as *Transient Traffic Flow*, is stored in the corresponding VOQ buffer of ONU_k and is retransmitted over the λ_{LAN} to reach its destination ONU, i.e. ONU_i. Finally, the destination ONU_i forwards the excess traffic to the destination eNB_i via the X2 interface. It must be noted that traffic redirection is typically performed for low-priority "besteffort" traffic (e.g., P2P file transfers, Web traffic, etc) appropriately marked by the AGW with the CoS value. Here, we propose a downstream wavelength sharing scheme, to support load balancing in the proposed WDM-PON ring architecture. The proposed scheme is implemented at the OLT, which has access to the traffic profile of all downstream wavelengths. It includes a periodic polling-cycle operation which is divided in two phases. In the first phase, overloaded and lightly-loaded wavelengths (or equivalently congested/uncongested ONUs) are identified. Further, excess traffic of congested ONUs and unused capacity of uncongested (or donor) ONUs is quantified. In the second phase, the λ_{LAN} channel is exploited to absorb temporal traffic overloads by redirecting traffic from congested ONUs through uncongested ones as pointed out in [9]. An algorithm is devised to dynamically match congested to uncongested ONUs. In what follows, we detail the two phases of the proposed scheme.

1) First Phase: Identification of Donor and Acceptor ONUs

In the first phase, the OLT periodically performs estimates of underlying traffic parameters and calculates the effective bandwidth C_{EB} of all downstream channels using the Norros formula [10]:

$$C_{EB} = \mu + \left[B^{H-1}k(H)\sqrt{-2\alpha\mu\ln e}\right]^{\frac{1}{H}} , \tag{1}$$

where $k(H) = H^{H}(1-H)^{1-H}$, B is the buffer size (in bits), μ the traffic mean rate (in bps), H the Hurst parameter of the traffic, and a the coefficient of variance. Effective bandwidth is a statistical estimate of the capacity required in order to satisfy a given QoS constraint (typically a buffer overflow probability). Lightly-loaded channels (which correspond to donor ONUs) are the ones with $C_{EB} < C$ and heavily-loaded ones (which correspond to congested ONUs) are those with $C_{EB} \geq C$, where C is the nominal wavelength capacity. Then, the OLT calculates a Traffic Counter (TC, in bps) for each ONU. The function of TC is to quantify the excess traffic or the unused capacity of congested or uncongested ONUs respectively.

For uncongested (donor) ONUs, the Traffic Counter (TC) corresponds to the unused downstream capacity. Assuming that N sub-flows are redirected through ONU_k and C_j the effective bandwidth of each flow, the following equation must hold so that the donor ONU_k is not overloaded:

$$\sum_{N} C_j + C_{EB}^{k} \leq C \tag{2}$$

This is a conservative estimate, as it does not account for statistical multiplexing gains from flow aggregation, providing a safety margin in case the underlying traffic profile changes. Thus, the donor capacity can be derived as the difference between the effective bandwidth of the downstream channel and the nominal wavelength capacity, i.e. $TC_k = C - C_{EB}^{k}$. Each time a new *Transient Flow* is redirected by a donor ONU, its TC is decremented accordingly.

For congested ONUs (where $C_{EB} \geq C$), TC corresponds to the effective bandwidth of the excess traffic component that must be redirected. For each congested ONU, downstream traffic is progressively split in a number of Transient Flows, until all excess traffic has been redirected. Each Transient Flow corresponds to a fraction of the overall mean rate, so that it can be matched and redirected over one donor ONU. Traffic splitting and redirection depends on the unused capacity of the associated donor ONUs (quantified with the Traffic Counter, TC) as well as the availability of the λ_{LAN}. The λ_{LAN} utilization profile is known to the OLT, since it is the OLT that arbitrates bandwidth allocations for the ONUs. Assuming ONU_i is a congested ONU, ONU_k is a donor ONU and $A_{k,i}$ the available bandwidth in the λ_{LAN} channel for the links in the path between ONU_k and ONU_i, the following equation must hold:

$$C_{i,k} = \min\left(TC_k, TC_i, A_{k,i}\right), \tag{3}$$

where $C_{i,k}$ is the effective bandwidth of the Transient Flow from (congested) ONU_i to (donor) ONU_k.

2) Second Phase: Dynamic traffic redirection

In the second phase, traffic redirection is implemented by setting up *transient flows* to be redirected from congested ONUs through uncongested ones, starting from the lowest priority classes. This requires a mechanism to efficiently split downstream traffic in multiple components (or sub-flows). In this work, a well-known stochastic hash-based technique is used for traffic splitting [11]. This technique involves calculating hash-based fingerprints on a packet-by-packet basis and classifying each packet to a sub-flow based on its fingerprint value. The fingerprint is calculated by applying a hash algorithm to the packet header 5-tuple, i.e., source and destination address, source and destination port, and protocol number. Fingerprint values are mapped to traffic sub-flows, that correspond to a fraction of the overall downstream traffic. For example, downstream traffic with mean rate m can be split into an excess traffic component, which corresponds to a fraction p of the overall traffic, with mean rate $m_{ex} = p \cdot m$. If no single donor ONU is able to accommodate the excess traffic component, the latter can be split in multiple sub-flows.

```
ForEach congested ONU_i
   ForEach upstream donor ONU_k
      C_i,k = min(TC_k, TC_i, Ak_i)
      If (C_i,k >0) Then
         Setup Transient Flow from ONU_k to ONU_i
         Update λ_LAN utilization profile
         Update TC_i, TC_k
         If (TC_i == 0) Then
            Break {no remaining excess traffic}
         EndIf
      Else {λ_LAN capacity exhausted}
       Report failure
       Break
      EndIf
   EndFor
EndFor
```

Figure 4: Pseudo code for Phase 2 of the bandwidth sharing scheme

At the beginning of the second phase, after acceptor/donor ONUs have been identified and excess/unused capacities have been quantified, the proposed scheme attempts to match donor and acceptor ONUs. It is evident that matching a congested ONU with the closest uncongested one(s) in the upstream direction minimizes the number of hops spanned by the transient traffic flows, which results in maximizing the spatial reusability gains of the λ_{LAN}. Thus, for each congested ONU, the OLT considers all upstream donor ONUs sequentially. Transient flows are setup from donor ONUs to acceptor ONUs until all excess bandwidth is redirected, or the λ_{LAN} capacity is exhausted. The basic algorithmic steps, which are repeated periodically at the OLT, are illustrated in Fig. 4. It must be noted that changes to the underlying traffic profile result in changes to the *transient flow* setup.

IV. PERFORMANCE EVALUATION

The proposed converged architecture along with the reservation protocol and dynamic bandwidth allocation / reallocation scheme were implemented in the ns-2 simulator framework and evaluated through extensive simulation experiments. Specifically, we simulated an LTE backhaul network with 8 ONU/eNBs, interconnected in a fiber ring with a 2 km diameter. An OLT was connected to the ring with a 20 km bidirectional feeder fiber. We modeled communication between each source-destination pair of ONUs (i.e. LAN traffic flowing through X2 interface) and between each OLT-ONU pair as separate traffic sources that generate packets according to a self-similar process, with H=0.7. Wavelength capacity was set to 1 Gbps, and buffer size of ONUs to 3MB. Finally, the DBA cycle time for LAN traffic t_{cycle} was set to 2ms.

In the first set of experiments, our objective was to evaluate the proposed ONU architecture, which supports all-optical meshing of ONUs using a dedicated λ_{LAN} channel. Without loss of generality, we limit this set of measurements to one pair of ONUs and assume symmetric upstream/downstream traffic load. In this simulation scenario, we implemented and compared the proposed ONU design with a simple reference design that does not incorporate the λ_{LAN} wavelength. In the reference design all inter-ONU traffic flows through the OLT, via the dedicated upstream/downstream wavelengths. Figure 5 displays the delay of the inter-ONU traffic (or LAN traffic) for both designs, when varying the LAN traffic load (denoted as ρ^{LAN}) and the upstream/downstream load (denoted as ρ^{UD}). Packet delay at the proposed ONU design is dominated by the queuing delay at the VOQs. At light loads the average queuing delay equals $t_{cycle}/2$, as packets wait for their designated timeslot. It can be seen that packet delay in the new architecture is stable and remains significantly lower than the packet delay of the reference architecture, especially at high loads. When the upstream / downstream channels are congested, packet delay at the reference architecture increases significantly.

Figure 6 displays the average delay of upsteam / downstream traffic vs. the traffic load, for different LAN loads. It must be noted that the average traffic delay of the proposed architecture remains unaffected by the LAN traffic, as the latter is transmitted in a separate channel. On the other hand, at the reference architecture LAN traffic wastes resources of the upstream/downstream channels, which leads to an increased delay. To this end, we may argue that the addition of a single wavelength, λ_{LAN}, significantly improves delay characteristics and offloads traffic from inter-eNB communication.

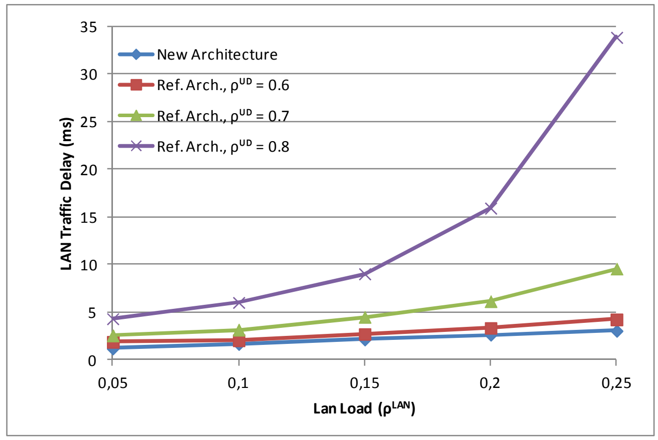

Figure 5: Inter-ONU (LAN) traffic delay vs LAN load for different upstream/downstream channel loads (marked ρ^{UD})

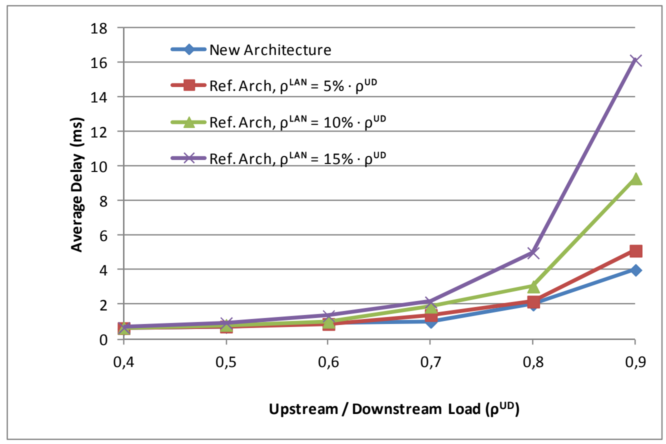

Figure 6: Average traffic delay vs. load, for different LAN traffic loads

In the second set of measurements, we evaluate the proposed scheme for dynamic bandwidth allocation and sharing of downstream wavelengths. This set involves communication among all ONUs that belong to the LTE backhaul. Among the 8 ONUs, 4 were randomly selected with input load $\geq$ 1.0 to simulate traffic overloads (congested ONUs) and 4 with input load $\leq$ 0.7 (uncongested ONUs). Figures 7 and 8 display the aggregated throughput of congested ONUs versus the downstream load. In Fig. 7 the load of uncongested ONUs (denoted with ρ) is varied, while LAN load is set to 0. In Figure 8 the LAN load is varied, while the load of uncongested ONUs is kept constant (ρ=0.1). For reference, the case of no redirection is shown in both figures.

It can be seen in both figures that the proposed architecture can effectively take advantage of unused capacity of uncongested ONUs as long as there is sufficient bandwidth availability in the λ_{LAN} channel. In particular, for the case of 1.8 load, the throughput of congested ONUs reaches 1.57 Gbps for uncongested ONU load of ρ=0.1 as can be seen in Fig. 7. This is a significant improvement, keeping in mind that the transient traffic would have been lost if traffic redirection was not implemented. As the uncongested ONU load increases to ρ=0.7 (and the unused capacity decreases accordingly) the congested ONU throughput gradually decreases down to 1Gbps. The same behavior is also recorded with the gradual increase of LAN load, as shown in Fig. 8.

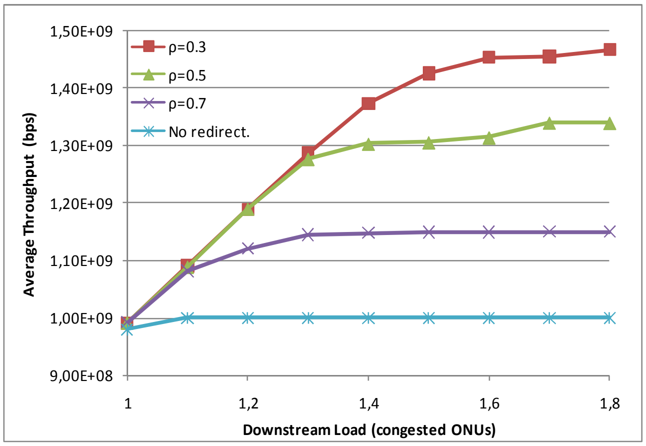

Figure 7: Average throughput of congested ONUs vs. downstream load. Uncongested ONU loads are denoted with ρ and LAN load is set to 0.

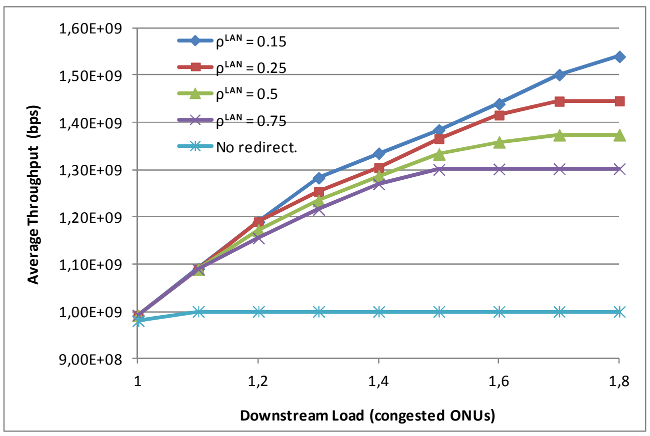

Figure 8: Average throughput of congested ONUs vs. downstream load for different LAN loads (marked ρ^{LAN}), and uncongested ONU load set to 0.1

Finally, Fig. 9 displays the end-to-end delay measured as the time that a packet leaves the OLT (or the originating ONU) until it reaches the destined congested ONU. In particular, Fig. 9 displays the delay of the LAN traffic as well as the delay of the downstream traffic (denoted as OLT traffic) for different loads of uncongested ONUs (ρ=0.3, 0.5, and 0.7), which are compared with the case of no redirection. From Fig. 9, it can be seen that the delay of the downstream traffic increases but never exceeds the case of no redirection. Thus, the QoS for delay sensitive services does not deteriorate when traffic redirection of excess traffic is employed. Furthermore, it can be seen that although the unused capacity of the λ_{LAN} channel is exploited for downstream traffic redirection, the delay of inter-ONU traffic remains unaffected and just marginally increases with the increase of the downstream load.

V. Conclusions

In this work a new unified PON-RAN architecture for LTE mobile backhaul networks was proposed, employing ring-based WDM PONs. The proposed architecture supports all-optical inter-BS communication as well as load balancing, via the dynamic re-allocation and sharing of downstream wavelengths. Lightpaths for inter-ONU communication are dynamically setup bypassing intermediate ONUs, thus avoiding unnecessary electronic processing. A dynamic bandwidth allocation scheme was also proposed for resource reservation and sharing of LAN traffic. It was shown through simulation that the proposed architecture significantly improves the delay characteristics of inter-eNB traffic, offloading upstream/downstream wavelengths. Finally, the proposed load balancing scheme effectively takes advantage of unused capacity, by redirecting traffic overloads to uncongested wavelengths.

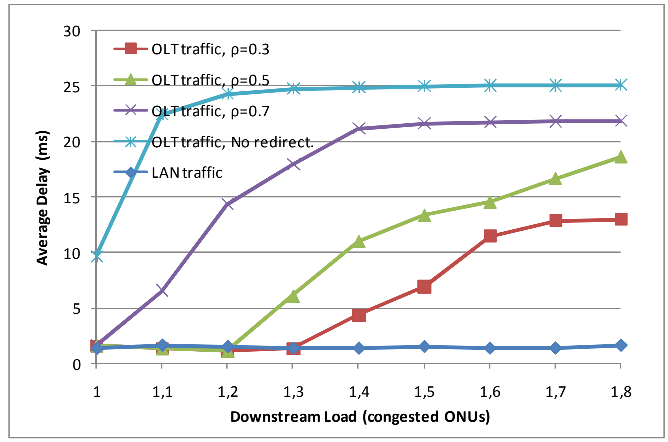

Figure 9: Downstream traffic and LAN traffic delay vs. downstream traffic load, for different uncongested ONU loads (marked ρ)

Acknowledgements

This work is supported by the Cyprus Research Promotion Foundation's Framework Programme for Research, Technological Development and Innovation 2009-2010 (DESMI 2009-2010), co-funded by the Republic of Cyprus and the European Regional Development Fund, and specifically under Grant ΤΠΕ/ΕΠΙΚΟΙ/0609(ΒΙΕ)/07.

References

[1] M.D. Andrade et al., “Evaluating Strategies for Evolution of Passive Optical Networks,” *IEEE Comm. Mag*., vol. 49, no. 7, pp. 176-184, 2011.

[2] E. Dahlman et al., *3G Evolution: HSPA and LTE for Mobile Broadband*, 2nd ed., Academic Press, 2008.

[3] D. Breuer et al “Opportunities for Next-Generation Optical Access”, *IEEE Comm. Magazine*, vol. 49, no. 2, pp. 16-24, 2011.

[4] S. Parkvall, A. Furuskar, and E. Dahlman, “Evolution of LTE toward IMT-Advanced”, *IEEE Comm. Magazine*, vol.49, no.2, 2011.

[5] M. Ali, G. Ellinas, H. Erkan, A. Hadjiantonis, and R. Dorsinville, “On the Vision of Complete Fixed-Mobile Convergence,” *IEEE/OSA Journal of Lightwave Technology*, vol.28, no.16, pp.2343-2357, 2010.

[6] C. Ranaweera, E. Wong, C. Lim, and A. Nirmalathas, “Next Generation Optical-wireless Converged Network Architectures”, *IEEE Network*, vol. 26, no. 2, pp. 22-27, 2012.

[7] K. Ramantas, K. Vlachos, G. Ellinas, and A. Hadjiantonis, “Efficient Resource Management via Dynamic Bandwidth Sharing in a WDM-PON Ring-based Architecture”, *Proc. of IEEE ICTON*, 2012.

[8] Y. Chen, C. Qiao, and X. Yu, “Optical Burst Switching: A New Area in Optical Networking Research,” *IEEE Network*, vol.18, no.3, pp. 16- 23, 2004.

[9] H. Erkan, ASM D. Hossain, R. Dorsinville, M. A. Ali, A. Hadjiantonis, G. Ellinas, and A. Khalil, “A Novel Ring-Based WDM-PON Access Architecture for the Efficient Utilization of Network Resources”, *Proc. of IEEE ICC*, pp. 5175–5181, 2008.

[10] I. Norros. “On the Use of Fractional Brownian Motion in the Theory of Connectionless Networks”, *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 6, pp. 953-962, 1995.

[11] Z. Cao, Z. Wang, and E. Zegura. “Performance of Hashing-based Schemes for Internet Load Balancing”, *Proc. of IEEE INFOCOM*, 2000.