UNIVERSITA' DEGLI STUDI
DI MODENA E REGGIO EMILIA

TCP Performance in Hybrid EPON/OBS Networks*

Maurizio Casoni, *Member, IEEE*
Department of Information Engineering - University of Modena and Reggio Emilia
Via Vignolese, 905 - 41100 - Modena - Italy
Tel.: +39-059-2056167, Fax.: +39-059-2056129, e-mail: maurizio.casoni@unimore.it

ABSTRACT

In this paper TCP performance are evaluated when high speed Ethernet over Passive Optical Networks are interconnected by means of a core optical network based on the Optical Burst Switching paradigm. The inter-working unit, or edge node, between these two networks is properly studied and discussed. A timer based assembly algorithm is here employed and it is analyzed and designed for operating with a EPON.
It is shown that different assembly time values have to be used for different TCP congestion window values, for optimizing the TCP throughput, and that longer assembly time values provide better fairness among TCP flows.

INTRODUCTION

The ever increasing popularity of hybrid networks, i.e. networks that have several different solutions and technologies, reveals that this kind of networks will play a crucial role in future infrastructures for multimedia applications.

Multimedia, interactive and high definition applications push for very high speed access network, where each user is given a bandwidth greater than 100 Mbit/s. Passive Optical Network (PON) is a promising technology to solve the last mile problem. Ethernet over Passive Optical Networks (EPON) has been regulated by IEEE and it is the high speed access network of this work.

Optical Burst Switching (OBS) can be seen as a middle term solution toward all optical packet switching whose goal is to improve wavelength utilization and sharing by introducing a dynamic wavelength management. In OBS networks data never leave the optical domain: for each data burst assembled at the network edge a reservation request is sent in advance as a separate control packet. There are two kind of nodes: edge and core routers. The main function of edge nodes is the burst assembly: they must collect IP datagrams and assembly them into burst according to proper assembly algorithms. Core nodes, on the other hand, deal with optical data bursts and the related control packets. They have to set up on the fly internal optical paths for switching bursts and take them hop-by-hop closer to their final destination. In addition, the offset time allows the core router to be buffer-less, avoiding then the employment of optical memories, e.g. fiber delay lines, required on the contrary by optical packet switching. The control packet carries relevant forwarding information, as the next hop, the burst length and the offset time. It precedes the data burst by a basic offset time, that is set to accommodate the non-zero electronic processing time inside the network, and it dynamically set up a wavelength path whenever large data flows are identified and need to traverse the network. Only the control packet is converted between optical and electronic domains, therefore is the only information delayed because of the conversion.

THE HYBRID NETWORK SCENARIO

The heterogeneous network here investigated is sketched in Figure 1. A EPON access network is linked to a optical backbone network which adopts the burst switching transfer mode. The inter-networking function is performed in the Inter-Working Unit (IWU) which is also the edge node of the OBS network. Each end-user, represented by a TCP New Reno agent with SACK option, is connected to a dedicated ONU and upstream transmissions from ONUs are regulated by OLT by means of the Interleaved Polling with Adaptive Cycle Time (IPACT) algorithm with the Limited discipline. As regards the IWU a timer based assembly algorithm is employed. Since the edge-to-edge delay has to be bounded, a maximum delay T_{max} can be tolerated in the assembly phase, but after that the burst must be transmitted. Given that IP packets within the same optical payload belong to the same class, two main different strategies can be adopted for upper layer information multiplexing, assuming information organized in flow as happens for TCP application: *per flow assembly*, where an optical burst contains data of the same flow and per-flow queuing is needed at the ingress of the assembly unit; *mixed flow assembly,* where a burst may contain information from different flows of the same class and per-QoS queuing is needed at the ingress of the assembly unit.

TCP performance are studied by evaluating two metrics, the throughput and the fairness. The throughput is a measure of the variability of the bandwidth usage over a given time scale. The average throughput is the amount of successfully transmitted bytes in a given time interval. As regards the fairness, when it is measured among flows of the same TCP flavor, it is referred to as *intra-fairness*.

Let B_{Pi} be the throughput of the i-th flow for a generic transport protocol P. Consider n flows employing the same protocol type P, and define $B_{Pmin} = \min[B_{Pi}]$, i=1,..,n and $B_{Pmax} = \max[B_{Pi}]$, i=1,..,n, the intra-fairness ratio for the considered streams is defined as $F_{intra} = B_{Pmin} / B_{Pmax}$ and the best intra-fair behavior implies $F_{intra} = 1$.

NUMERICAL RESULTS

In this section, numerical results on TCP performance regarding the reference scenario are reported. They have been obtained by means of simulations using the ns-2 simulator vers. 2.31. Users are represented by N=10 TCP with SACK option agents and they are connected to a dedicated ONU. CBR is the application over the TCP agents on the customer side. OLT is to its turn connected to the OBS edge node with a 1 Gbit/s link with 1 ms delay, representative of a 200 km distance. The maximum cycle time *Tc* has been considered in the range of (0.5-20) ms. In the edge node a mixed-flow assembly algorithm is assumed with T_{max} in the range of (0.5-4) ms. The OBS network is approximated as a link operating at 2.5 Gbit/s with burst loss probability $P_b=10^{-3}$. This link is assumed to introduce a delay of 5 ms, representative of the propagation delay over a distance of 1000 km. A egress edge processes burst and delivers datagrams to TCP sinks. Figure 2 shows the average TCP throughput as a function of T_{max} and different values of the congestion window. It is evident to observe that the throughput remarkably improves by increasing the congestion window from 64 to 256 segments and this is due to the fact that each TCP source can send more segments during its time interval *Tc/N*. Even more interesting is that the performance have different point of maximum for the three curves. As regards fairness, Table 1 reports the intra-fairness index for different values of the maximum congestion window and T_{max}. The basic outcome here is that for longer assembly times the intra-fairness remarkably improves for all values of the congestion window.

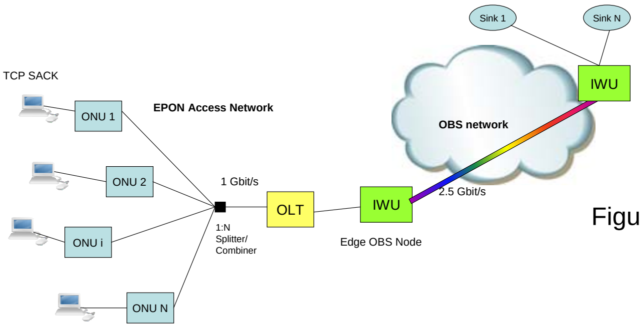

Figure 1

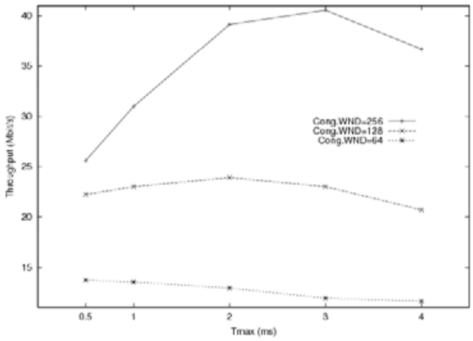

Figure 2

T_{max} *cwnd*	0.5 ms	2 ms	4 ms
64	0.968	0.999	1
128	0.93	0.94	0.9995
256	0.883	0.95	0.9995

Table 1

*The work described in this paper was carried out with the support of the BONE-project (Building the Future Optical Network in Europe), a Network of Excellence funded by the European Commission through the 7th ICT-Framework Programme.

UNIVERSITA' DEGLI STUDI
DI MODENA E REGGIO EMILIA
Dipartimento di Ingegneria
dell'Informazione

Research Group
Wireless & Optical Networks and
InterNEt Technologies

Prof. Maurizio CASONI
E-mail: maurizio.casoni@unimore.it
http://www.dii.unimo.it/casoni