OBiS: Optical Bit Switcing – A Switch Architecture for Optical Omnipresent Ethernet

Ashwin Gumaste, Saurabh Mehta, Santosh Rana and Tamal Das
Department of CSE, Indian Institute of Technology, Bombay, India. 400076.
Email: ashwing@ieee.org, saurabhm@iitb.ac.in, skrana@cse.iitb.ac.in, tamal@cse.iitb.ac.in

Abstract: The Omnipresent Ethernet architecture has been proposed for end-to-end communication using pure Ethernet. We extend the scheme to an all-optical variant that supports a novel scheme called optical bit-switching which is verified through simulations.

OCIS codes: 060.4250 Networks; 060.4250 Networks

1. Introduction

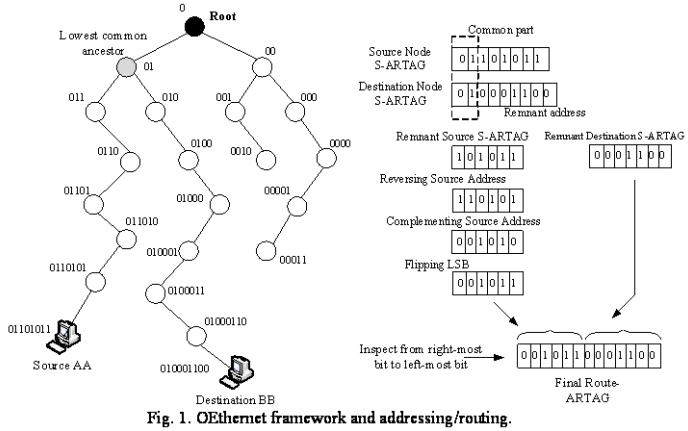

Fig. 1. OEthernet framework and addressing/routing.

Ethernet has traversed an interesting journey from its inception to the present day from being a CSMA oriented LAN technology to switched Ethernet at Gigabit speeds, to contemporary Carrier Ethernet that guarantees deterministic delay and supports service attributes. We have used different aspects of Ethernet technology to propose an end-to-end all-Ethernet solution called *Omnipresent Ethernet* (OEthernet) that provides communication with and without IP, abstracting higher layer functions to the lower layers in an effective manner. Multiple aspects of OEthernet have been studied in [1-2]. In this paper, we propose an all-optical implementation of OEthernet using a paradigm called OBiS or Optical Bit Switching. We begin by first recapping OEthernet and then showcasing the technological advances of OBiS. If we consider the interconnection model in networks today (for example [5]), then we observe that at a logical level, all networks essentially lead to a hierarchical tree, though physically, these might be stars, rings or mesh. A hierarchical tree is typically characterized by a random distribution of degrees of connectivity in its nodes. We can *smoothen* the degree of connectivity by adding virtual nodes after which every node in the network is characterized as a binary node (having connectivity 1x2 or 1x1). The process of smoothening is shown in [1-2] and leads to the network being abstracted to a binary tree (with all nodes being binary nodes). The advantage of having binary nodes is that we can adapt binary routing and source routing, which simplifies the forwarding plane. A node is allocated a binary address (called *source address-route tag* or S-ARTAG) which essentially is a route from the node to the root (of the binary tree). To derive this address we traverse the path from the root to the node, and for every right turn that we take, we add to the existing address by a 0 bit, while for every left turn that we take, we add to the existing the address by a 1 bit. In this way a binary address is created. A source node can communicate to a destination node by creating a binary string (called route ARTAG or R-ARTAG) that gives the address from the source to the destination. To do so, the source node takes the S-ARTAG of itself and that of the destination, compares the two (by matching the MSB), discards the leading common bits, does a 1s complement on the source remnant string, does a further 1s complement on the LSB of the source remnant string and then conjoins the two remnant strings. This procedure is shown in Fig. 1 for node AA to communicate to node BB. The benefits of source and binary routing are tremendous – high-reliability, low end-to-end delay, lower energy use and others as mentioned in [1]. To implement such a scheme in a network, we choose a modified Carrier Ethernet implementation mechanism. In particular, we switch off spanning tree and add additional VLAN tags just like in the PBB-TE or MPLS-TP drafts, but with a twist! We insert the S-ARTAG and R-ARTAGs as VLAN tags (hence the name ARTAG) in the Ethernet frame. Additional VLAN tags are also possible as part of the OEthernet framework, as shown in Fig. 2. The functioning of OEthernet is discussed in [1] and the key aspect of the

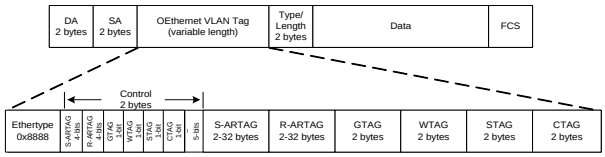

Fig. 1. OEthernet frame structure.

forwarding plane is the fast lookup at every node – which requires the isolation of just one-bit. Switching at a node hence occurs based on whether this bit is a 0 or a 1. It is well known that there is an advantage of switching in the optical domain (like Optical Packet Switching), in terms of cost, speed and energy consumption. This paper proposes a switch architecture and a method to *implement the OEthernet scheme in the optical domain* based on a concept we call as OBiS or optical bit switching. Section 2 describes how OBiS can be implemented in the OEthernet framework. Section 3 discusses simulation results, while Section 4 concludes the paper.

2. OBiS: Optical Bit Switching – Technology Enabler for the Optical OEthernet Framework

We will demonstrate communication in the binary tree using all-optical technology in conjunction with OEthernet concepts. We assume that the destination binary address is known in prior to the source node, though the address-fetching task can also be simplified and adapted in its entirety at the optical layer. Consider the binary tree as shown in Fig. 1. We will now consider node 01101011, 0100011000 as two nodes for case study and show how we can implement OEthernet in the all-optical set up. The architecture for all optical OEthernet solution is modified from the traditional OEthernet switch using Sagnac loop technology [3] that provides the basic building block for all-optical switching. The idea is to switch a frame between the cross port and the bar port of a 2x2 switch by making use of the value of a single bit in the optical domain. The challenge then is to:

(1) Detect an arbitrarily long incoming frame in the optical domain, using an optical delimiter technology that allows us to indicate the *start of a frame*.

(2) Locate the corresponding bit in the frame that would indicate to the switch whether the frame should be switched to the cross port or to the bar port. This bit is defined as the *node decision indicator* (NDI) bit.

(3) Based on the NDI bit logic-value, perform switching while keeping the bit and frame in the optical domain.

Assumptions: To perform all-optical switching we assume that the gap between 2 frames is exactly an integer multiple of IPG (*Inter-Packet Gap*) which we assume to be 64 microseconds. We also assume that the line-rate is fixed at either 1 or 10 Gbps without line encoding (8b/10b). The frame format is the same as shown in Fig. 2, except that by default the first tag amongst all the VLAN tags is the R-ARTAG.

Architecture: The system works as follows: initially there is a manual boot up phase in which every 1x2 switch in the binary tree is fed its NDI number. The network administrator manually computes the NDI number for every node – which is a unique number (for that level in a binary tree). The NDI number is in fact the argument value of the R-ARTAG. The NDI number tells which bit from the 1st bit in the R-ARTAG, is the NDI value.

A frame arrives at a node and it goes through an IPG+SOF (Start of Frame) sub-system (see Fig. 3). This sub-system consists of an optical coupler and a delay line. The coupler has a 99:1 power splitting ratio, while the Fiber Delay Line (FDL) delays the incoming frame by using signal recirculation for about 10 microseconds – enough time for the tapped 1% power to be detected and a decision to be made in the electronic domain, as to whether the incoming data stream is an OEthernet frame (from the SOF sequence detection). The OEthernet frame then moves to the "decision box" which has the task of extracting optically the NDI value from the frame. The NDI is electronically extracted by the IPG+SOF block.

The decision box consists of an FDL of duration $(SOF+NDI)*T_b+T_p$ where SOF is the number of bits in the start of frame field, T_b is the bit duration; T_p is the processing duration of the Sagnac loop (switching delay). After SOF+NDI duration of time has elapsed since the first bit of the frame entered the FDL, the NDI arrives at the entrance of the FDL. At that time instance, the coupler, at the entrance of the FDL couples 2 % of the optical power to a drop port. To do so, this coupler is exemplified by a fast moving (acousto-optic or electro-absorption based) Mach Zehneder Interferometer (MZI) based device that can switch within T_p+T_g duration. The frame (while the switch is being configured), is delayed in an additional FDL. To compute the exact time when this power splitting is done, an electronic counter is instantiated to drive the MZI circuit – such that the counter resets upon the arrival of a new frame (when no information arrives for at least IPG duration). The MZI now couples the power for T_b-T_g duration, where T_g is the system guard-band. The coupler has now extracted an optical copy of the NDI and this feeds to the control line that drives the Sagnac loop. The Sagnac loop in Fig. 3 consists of a fiber connected in a loop to each end of the 2x2 coupler (50:50 splitting ratio). Control pulses are fed into this loop by another 3dB coupler. In order to reduce the switching power and create high nonlinearity in the loop, a Non Linear Element (NLE) is present, which in our case is a 12 dB gain flattened Semiconductor Optical Amplifier (SOA).

Working: Consider a case when there is no control pulse (NDI value = logical 0) and an incoming frame arrives at the base 2x2 coupler. The frame is optically split into two counter propagating signals which interfere with each other at the midpoint of the loop and then move away toward the base coupler. Since, power of the two counter propagating waves is low, no significant XPM effects are observed. The NLE (SOA) exhibits transparent response and does not affect the two interacting waves. The two copies of the frame again interfere at the base coupler and are

reflected back with zero transmittivity. The control signal (NDI bit) is a high power signal (after passing through the EDFA) and hence creates non-linearity in the SOA. The non-linearity rise time is less than the data bit time and fall time is slower than a multiple of bit times – T_b. Now, we consider a case, where the control pulse (logical value 1) is launched such that it crosses SOA, just after data pulse-train, which is traveling clockwise. When the counter clockwise pulse reaches the SOA, it encounters a change in the refractive index and undergoes phase change with respect to the clockwise traveling pulse. These two pulses then interfere at the base coupler and the data pulse-train now gets transmitted. Hence, in summary, to achieve 1 x 2 switch operation, we *apply no* control pulse (logical 0) to the SAGNAC loop when we want to switch from input to bar port and *apply a* control pulse (logical 1) when we want to switch from input to cross port as shown in Fig. 3.

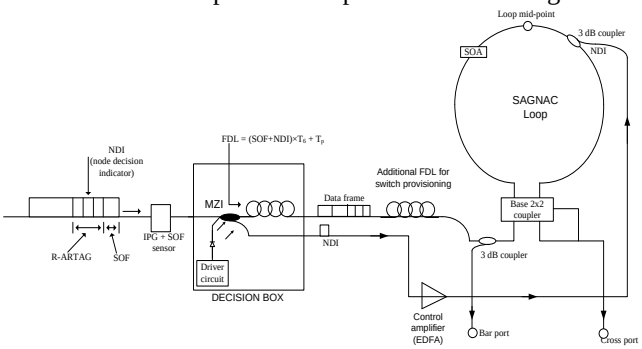

Fig. 3. OEthernet 1x2 switch implemented in all-optical technology.

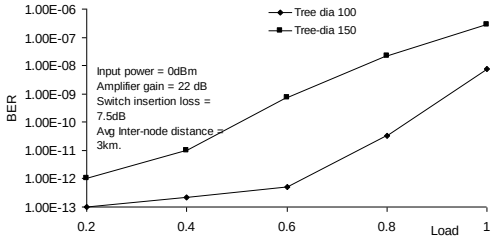

Fig. 4. BER versus Load for two binary trees.

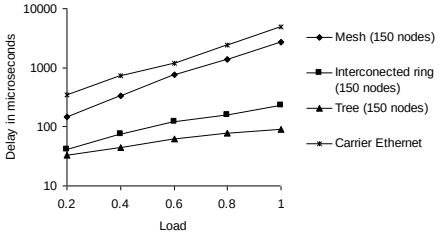

Fig. 5. Delay as a function of load.

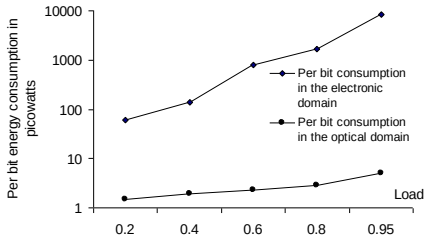

Fig. 6. Energy consumption of OBiS versus Carrier Ethernet.

In this way the NDI facilitates switching between the bar and cross ports of the Sagnac loop switch resulting in all optical Omnipresent Ethernet. Shown in Fig. 4 is a view graph that shows the performance of the all-optical OEthernet switch. To get the results in the figure we simulated the switch in Fig. 3 using a Matlab code. The system parameters are embedded in Fig. 4 (input power =0dBm, EDFA gain = 22dB, insertion loss = 7.5 dB). Two kinds of trees were simulated – one with a diameter of 100 nodes (from root to leaf) and another with diameter of 150 switches. Results show good performance for both the schemes – with BER always better than 10E-8 even for loads around 80 %. Shown in Fig. 5 is a viewgraph of delay as a function of system load (defined as the ratio of the total data entering the system to the maximum possible system handling capacity). We simulated a mesh network of 150 nodes, a 150-node interconnected ring network (with each ring of 10 nodes), a tree of 150 nodes and a tree of 150 nodes but having Carrier Ethernet (MPLS-TP format). As can be seen, the delay for tree Optical OEthernet is the lowest and is almost two orders of magnitude lower than the MPLS-TP scheme. Shown in Fig. 6 is a viewgraph of per-bit energy consumption of the OEthernet framework with OBiS technology and conventional Carrier Ethernet implemented by electronic switching technology. Commercial Carrier Ethernet switches as shown in [4] have been assumed. The OBiS scheme has on an average two orders of energy efficiency over the Carrier Ethernet scheme.

3. Conclusion

We have presented an all-optical implementation of the OEthernet concept that leads to low-latency and low-energy consumption in core networks. The OEthernet concept has significant performance improvement over conventional networks, and this is only enhanced with the addition of the O-BiS technology solution.

References

[1] A. Gumaste et. al, IEEE/OSA Journal of Lightwave Technology, 2010.
[2] S. Mehta, R. Vaishampayan, A. Gumaste, Optical Fiber Commun. Conf. Post Deadline paper, PDP4 San Diego, USA.
[3] J. P. Sokoloff, P.R. Prucnal, I. Glesk and M. Kane, *IEEE Photon. Technol. Lett.*, 5, 787-790 (1993).
[4] Online: http://www.cisco.com/en/US/products/hw/switches/ps637/index.html
[5] Online: http://en.wikipedia.org/wiki/Internet.