Achieving Multi-Rate Dynamic Sub-Wavelength Service Provisioning in Strongly Connected Light-trails (SLiTs)

[1]Ashwin Gumaste, [2]Nasir Ghani, [1]Paresh Bafna, [1]Akhil Lodha, [1]Sumit Srivastava,
[1]Tamal Das and [3]Si Qing Zheng
[1] Depts of CSE/EE Indian Institute of Technology, Bombay, [2]Tenessee Tech University, Cookeville USA, [3]Univ. of Texas, Dallas
Email: ashwing@ieee.org, nghani@tntech.edu, {paresh@it, alodha@cse, sumits}.iitb.ac.in, tamal.das@iitd.ac.in, sizheng@utdallas.edu

Abstract: We report on achieving multi-rate, dynamic, sub-wavelength provisioning of VoIP, video, storage and data services over Strongly connected Light-trails (SLiTs). A novel service differentiation subsystem and associated control are discussed and test-bed results are presented.

OCIS codes: 060.4250 Networks; 060.4250 Networks

1. Introduction

The shift of revenues from traditional voice services to a multitude of VoIP, *Video-on-Demand* (VoD), *Pseudo Wire Edge-to-Edge Emulation* (PWE3), Triple play and Data-Centers is a strong motivation for new low-cost and dynamic optical layer solutions in metro environments. Conventional SONET/SDH hierarchy is now being replaced by more data-centric packet aware technologies like GigE lightpaths and *Resilient Packet Rings* (RPR) with IP/MPLS overlay. GigE is not efficient (on account of its requirement of end-to-end wavelength granularity) nor is it dynamic, while RPR is expensive (due to OE and EO conversions at every node). A new approach is required that provides efficient grooming of sub-wavelength traffic preferably at the optical-layer while enabling necessary dynamic bandwidth provisioning thus facilitating emerging services. We report a solution that enables sub-wavelength all-optical (spatial) grooming, multi-service support at multiple line-rates. The proposed solution is deployed through a metro ring WDM test-bed called DynaSPOT – Dynamic Service Provisioned Optical Transport, and is built using Strongly connected Light-trail (SLiT) technology [1, 5]. A light-trail [2] is a unidirectional optical bus that is provisioned through an Out-Of-Band (OOB) control channel facilitating spatial sub-wavelength grooming of traffic along multiple nodes. A SLiT [1] is a bidirectional implementation of a light-trail with the node architecture shown in [1] and bidirectional EDFAs for implementation in metro rings. SLiT bandwidth is arbitrated by a *controller* and nodes time-share the bandwidth to provision *connections*. Connections are provisioned using burst-mode optics (like in light-trail [2]). In the DynaSPOT test-bed we show how to dynamically provision these connections with a particular emphasis on services. The time-penalty incurred due to queueing of data while facilitating time-sharing is made proportional to the particular service latency type by a new transponder subsystem. Another salient feature of the DynaSPOT test-bed is the ability to use the same SLiT for connections at multiple line-rates, doing so in a dynamic fashion. This feature enables nodes with diverse transponder inventory to use the same SLiT, thus reducing capital expenditure and maximizing efficiency. A final focus of the DynaSPOT test-bed is on data-center networks and the ability to lead to optical layer driven virtualization and consolidation features.

2. Description of the DynaSPOT (Dynamic Service Provisioned Optical Transport) Test-bed

The DynaSPOT test-bed is a 4-node open-optical metro-ring that can support single wavelength communication at both 1 Gbps and 2.5 Gbps. The following design choices are involved in the DynaSPOT test-bed:
SLiT: A four-node SLiT is created to facilitate sub-wavelength dynamic service provisioning. A node can communicate to any of the other 3 nodes using SLiT principles of communication (time-shared) mentioned in [1]. An out-of-band control channel is used for arbitration. Each node is provided with a PowerPC embedded in an FPGA (*Virtex 2Pro*) for arbitration and control purposes. Nodes can establish connections with other nodes that have compatible receivers – i.e. a source node can communicate with destination node(s) under the condition that the destination node has a receiver at the same line-rate as the source. In this way multiple nodes can time-share the SLiT at different line-rates. For example, in the 4-node SLiT, nodes N_1 and N_3 communicate at 2.5 Gbps while nodes N_2 and N_4 communicate at 1 Gbps. Further, node N_1 can receive information at both 2.5 Gbps as well as 1 Gbps. The SLiT is assumed to be time-slotted with slots of 400 *us* duration separated by guard-bands of 20 *us*.
Node architecture: The node architecture used in DynaSPOT is shown in Fig. 1, a critical subsystem called *trailponder* is shown in Fig. 2, and the conceptual layout as shown in Fig. 3 and photographs of the test-bed are shown in Fig. 4. The node architecture has evolved from a ROADM, with incoming WDM signal de-multiplexed by an AWG (Arrayed Wave Guide). Constituent wavelengths (ITU defined in C-band) are fed to a SORS (SLiT Optical

Retrieval Section) that consists of a series of two ON/OFF (optical) switches and two passive optical couplers in 2x1 configuration. Incoming signal (in either direction) is dropped and continued at the *drop-coupler* while local signal can be added into the SLiT passively through the *add-coupler*. The two couplers have 50/50 splitting/combining ratio and the two switches are in the ON state at all intermediate nodes in the SLiT. In addition, a single VOA is used to stabilize optical power-level. To add and drop signal, to and from the time-shared SLiT as well as to facilitate service provisioning, we use the trailponder. The trailponder (as shown in Fig. 2) consists of burst-mode optics (laser and receiver) [3] and associated electronics (memory, processor) and is triggered through a control card. The trailponder is analogous to a transponder – it facilitates client signals to be transmitted (and received) over the SLiT. The OE (and EO) trailponder card has the added function of storing data as well as scheduling stored data in an optimal manner to facilitate service provisioning. To do so, the trailponder uses either a VLAN based differentiator that segregates incoming packets (from a client) based on VLAN tags (explained later), or uses the *type* field for differentiation if the transmission is based on Generic Framing Protocol (GFP). Packets based on service types are stored in corresponding *service buffers*. Whenever a node is granted a slot for transmission, buffers are emptied in the following priority order: storage→voice→video→data. The total memory allocated for buffering in a trailponder is 2Mb (1Mb for TX, 1Mb for RX). Further, the 4 service buffers are of size 200, 200, 200 and 400 kb with the largest memory block allocated for data traffic.

Control card is built using a Virtex2 FPGA board (XUP V2P) and is connected to all the nodes through a 10/100 Ethernet switch. For sake of simplicity it is assumed that the control channel is non-blocking, implying that when two nodes send requesting signals on the control channel the two signals (in form of Ethernet packets) do not collide.

Connection provisioning: a node requests a slot to an arbiter node (node N_3 in our case). The request is sent as a *utility valuation* (explained later). A node sends a request (through the control channel) during every slot for bandwidth provisioning in the next slot. The arbiter decides which node would transmit in the next slot based on the highest utility valuation and this node is *granted* permission to set up a connection (max duration 400 *us*).

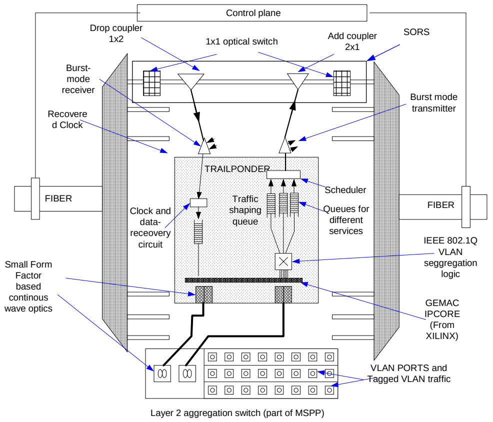

Fig. 1. Node architecture and conceptual SLiT model

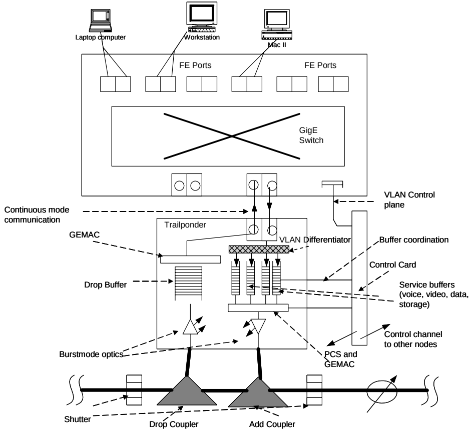

Fig. 2. Trailponder (for VLAN support) and control card

VLAN/GFP type based service differentiation: As shown in Fig. 2 traffic from clients arrives into the trailponder in a continuous-mode fashion (typically through 10/100/1000 Mbps interfaces). This traffic is aggregated by a layer-2 switch and then sent to the trailponder. The trailponder receives layer-2 frames (Ethernet) with VLAN (IEEE 802.1Q) tags and either sends them into the network as Ethernet frames (at 1 Gbps), or maps these into GFP frames (at 2.5 Gbps), while preserving mapping between ingress VLAN tag and egress GFP-type field. These frames (Ethernet/GFP) are queued in service buffers. For the former case of end-to-end Ethernet traffic, the frames enter the trailponder FPGA through a Xilinx GEMAC, that is engineered to capture VLAN tags as well as to perform PCS (Physical Coding Sub-Layer) functions. Since VLAN tags have 3-bits allocated for service type, the FPGA is able to store the incoming data (within the Ethernet frame) into one of the 4 appropriate buffers (as shown in Fig. 2). The information is stored into a buffer depending on the match of the buffer id (1=voice, 2=video, 3=storage and all other=data) to the corresponding VLAN tag or GFP-type.

Utility Valuation and Provisioning: The trailponder computes a value called *time-to-service* for a buffer, defined as the time remaining before which a buffer must be scheduled (into the SLiT) or else the longest waiting packet (of that service) would be timed-out (service latency not met). Similarly the trailponder also computes a value of *buffer utilization*, defined as the ratio of number of bits in the buffer to the total buffer capacity. From time-to-service the trailponder computes *service valuation* [2] as:= *1/{1+time-to-service}* for every buffer. Since both service valuation and buffer utilizations are entities yielding ratios in the range [0,1] the trailponder passes on the maximum of the two ratios to the control card. Each control card then sends a utility valuation to the arbiter. The utility valuation is the maximum value over all buffer utilizations and service valuations. The control card also sends a *destination* list to the arbiter – that consists of all possible destination MAC addresses of packets stored in the buffer. The multiplicity of destinations facilitates optical multicasting feature in the SLiT bus.

Multi-Speed support: When a connection is provisioned at a certain line-rate the source and destination node are assumed to have the requisite laser/receiver at that line rate. It may however happen that a node has 2 receivers at different line rates (1 Gbps and 2.5 Gbps) connected passively (in drop and continue). In such a case, the node must switch OFF the receiver that is not in sync with the line rate of the connection. This is done ahead in time through the OOB control packet (*grant*) that is sent by the arbiter node and triggers the receiver bias OFF.

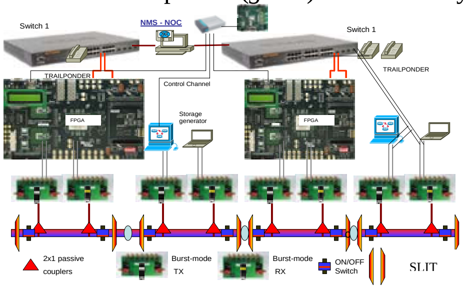

Fig. 3. Conceptual layout of the test-bed.

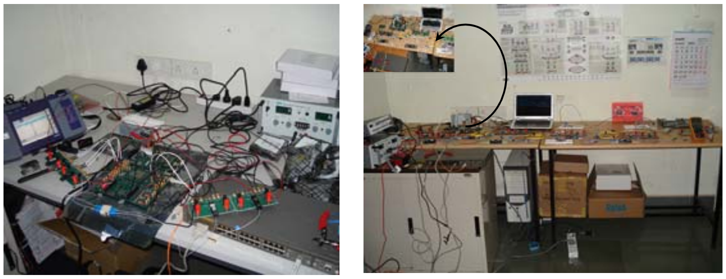

Fig. 4. Test-bed pictures (single node, 4-node configurations).

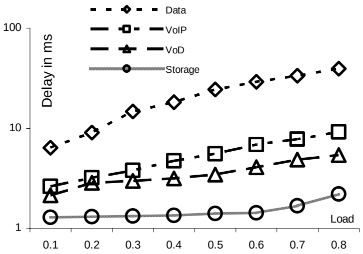

Fig. 5. Delay as a function of load.

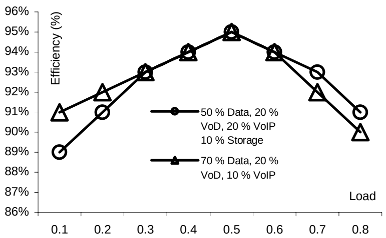

Fig. 6. Efficiency of the system.

3. Experiment and Results

We performed an experiment with 4-node SLiT at 1550 nm. The SLiT supported line-rates of 1 Gbps (GigE) as well as 2.5 Gbps (GFP). The setup assumed four traffic types –VoIP, VoD, data and storage; feeding into an aggregation switch at all the nodes. All Ethernet traffic was appended with VLAN tags while GFP traffic was differentiated based on type field. Traffic was varied by a generator through a layer-2 aggregation switch as well as a Gigabit Ethernet VLAN tester that generated tagged VLAN based packets. 2 nodes were connected to the aggregation switch while 2 other nodes were connected to the Gigabit Ethernet VLAN tester. Traffic intensity (load) could be varied at both generators. Measurements were performed by increasing traffic from 100 Mbps to 800 Mbps for GigE and 200 Mbps to 2 Gbps for GFP. Shown in Fig. 5 is the average end-to-end delay profile for data, VoIP, VoD and storage traffic as a function of load while shown in Fig. 6 is the efficiency of the system with two different traffic mixes. In Fig. 5 we observe that the average end-to-end delay is well within the acceptable service latency requirements even when the SLiT is heavily loaded with duplex VoIP traffic and sensitive storage traffic. In Fig. 6 we observe that with more data traffic (and less dynamic – storage traffic), efficiency is better at lower and medium loads but degrades at heavy loads due to data-burstiness. We also observe that storage traffic (with dynamism for virtualization/consolidation) is efficiently provisioned through the SLiT and associated dynamic control.

4. Conclusion

We demonstrate the DynaSPOT test-bed to support sub-wavelength grooming, dynamic service provisioning (VoIP, data, VoD and storage) and multi-rate communication over a single wavelength using SLiT technology. An architectural overview and experimental results are presented.

Acknowledgement: Authors thank TCS-TRDDC for financial support, R. Nachane (JDS Uniphase), and R. Subramaniam, N. Varma (Xilinx/Avnet) for their equipment support and Prof. Krithi Ramamritham (IIT Bombay) for his encouragement.

5. References

[1] A. Gumaste, S. Jain, N. Ghani and S. Zheng, *IEEE/OSA OFC 2006 Anaheim CA 2006.*
[2] A. Gumaste and S. Zheng, *IEEE Communications Magazine Mar. 2005. Vol. 21 No. 3 pp 72-78.*
[3] Y. Ota and R. Swartz, *IEEE Journ. of Lightwave Tech. 1990 Vol. 8 No 12.*
[4] A. Gumaste et al, *IEEE ICC 2007* June 2007, *Glasgow UK.*
[5] P. Bafna, A. Gumaste and N. Ghani, *IEEE/OSA OFC 2007, Anaheim CA TuG.*