Providing Ethernet Circuit Quality of Service and High Bandwidth Efficiency through Fusion Networking

Raimena Veisllari[1], Steinar Bjornstad[1,2], Kurosh Bozorgebrahimi[3], and Norvald Stol[1]

[1]Department of Telematics,
Norwegian University of Science and Technology, Trondheim, Norway
{veisllar, norvald.stol}@item.ntnu.no

[2]TransPacket, Drammensveien 134, 0277 Oslo, Norway
steinar@transpacket.com

[3]UNINETT, Abels gate 5, 7465, Trondheim Norway
kurosh@uninett.no

Abstract

In this paper we propose the Fusion network as a solution technology that enables the deployment of Ethernet in carrier and mobile-backhaul networks by adding a circuit transport service while maintaining Ethernet's important property of low-cost high-throughput. We experimentally demonstrate, through a field-trial in the UNINETT carrier network with real production traffic, that the Fusion network enables circuit quality transport on an Ethernet wavelength: low deterministic delay, delay-variation in the range of 15 ns and no packet loss. The high resource utilization is enabled by the insertion of statistically multiplexed (SM) traffic on the same wavelength dedicated to the circuit service; SM traffic is added transparently and without affecting in any way the circuit service. Results show that the resource utilization is increased to an average of 88% by inserting 78.7% additional Ethernet packet-traffic on the production traffic's wavelength link.

1. Introduction

An important network evolution has been taking place in the last decade: the migration from circuit-switched towards packet-switched technologies in the core part of the network. A key motive has been the ever-increasing bandwidth demand and change in the traffic patterns. Packet-based networks meet the high-throughput requirements with lower cost technologies by employing statistical multiplexing that maximizes the utilization of the network resources. In this context, service providers and the research community have seen the potential for growth and revenue of Ethernet services in carrier networks [1], [2]. Hence, significant effort has been dedicated to developing and standardizing carrier Ethernet solutions. Nevertheless, existing legacy circuit-switched technologies are still widely deployed and a considerable source of revenue; most importantly the services they have been providing must be retained. For example, replacing leased line technologies means that the service they provide should still be provided to the customers that use it, with the same quality of service (QoS). From a customers' perspective, it is crucial that the network is able to provide with the same service level agreement (SLA) the same class of service, e.g. real-time applications like high-quality video-conferencing that require low delay and jitter. Thus, carrier-grade Ethernet should not only interwork with existing legacy technology as it migrates to a

This paper was presented at the NIK 2013 conference. For more information see http://www.nik.no/

carrier network [3] but moreover, once there, it should be able to provide as a minimum the same range of services that it is replacing.

Statistical multiplexing, which offers the main benefit of exploiting the network capacity efficiently, also imposes a crucial drawback: Ethernet networks are high bandwidth-delay product networks. Delay and packet delay variations (PDV) are imposed on the traffic because buffering is needed for handling the statistical variations in the traffic pattern. While the high-throughput is welcomed by bandwidth hungry applications and data centric technologies, the delay and PDV characteristic is not well-fitted for the transport of time-sensitive information.

An important example is the deployment of carrier Ethernet in the mobile back-haul network [4], [5]. Mobile base-stations require synchronization signals for coordinating hand-over. The replacement of circuit-switched SDH/SONET, that inherently carries the time information, imposes the need for transporting synchronization information at the packet-layer across the asynchronous Ethernet network. Furthermore, packet-layer synchronization solutions using e.g. the IEEE 1588 Precision Time Protocol (PTP), rely heavily on a low network packet delay variation (PDV) for achieving high precision [6].

Priority scheduling with efficient tag-based switching [7] or Ethernet transport level congestion mechanisms [8] have been introduced for differentiated QoS profiles. However they cannot provide the hard QoS of a dedicated circuit, e.g. a dedicated wavelength lightpath in an optical wavelength routed optical network. Emulations of leased lines through the E-Line service of Ethernet virtual connections have been specified for this purpose [9]. The Ethernet Virtual Private Line (EVPL) E-Line solution still employs statistical multiplexing in sharing the same physical resource between multiple EVPLs. The Ethernet private line (EVP) E-Line solution does not apply statistical multiplexing but dedicates the whole physical resource to the service. The latter, e.g. Ethernet over dedicated wavelengths, enables Ethernet circuits with zero packet loss, low delay and minimized PDV end-to-end across the network. However, the drawback is that by fully reserving the wavelength to a service, it is not possible to exploit the vacant leftover capacity. Furthermore, the intermediate transit nodes in the network cannot increase the resource utilization as statistical multiplexing is no longer applied.

In this paper we propose the Fusion network as a solution technology that enables both Ethernet transport with circuit QoS and increased resource utilization by means of statistical multiplexing. Fusion networking builds on the integrated hybrid optical network (IHON) architecture [10-12], but moves the architecture from all-optical switching technologies e.g. optical packet switching, towards using standard Ethernet technology over the optical medium. The Ethernet packet traffic is classified into two classes: guaranteed service transport (GST) which is given circuit service with absolute priority and statistical multiplexed (SM) lower priority traffic; both are fused in a time-interleaved manner in the same physical resource without using time-slots. The GST stream is provisioned with an end-to-end dedicated lightpath in the network, i.e. there are no losses or collisions with other GST streams. An SM packet scheduler looks into the GST stream of packets passing through the link, detecting vacant time gaps when the channel is idle. Whenever a gap is detected, the scheduler searches the input queues for an available SM packet of suitable size that fits the gap. If so, the packet is inserted in the gap without affecting the timing of the packets in the GST stream. The result is a network with a unique combination of properties: (1) GST circuit paths with low processing overhead, zero packet loss, low delay and minimized PDV combined with (2) packet paths with SM capability enabling the throughput efficiency of packet networks. We experimentally demonstrate the circuit QoS and high throughput efficiency through field-trial results in the carrier network of UNINETT with Fusion prototype nodes from TransPacket.

The rest of this paper is organized as follows. In Section 2 we introduce the main characteristics of Fusion networks and the hybrid node architecture. In Section 3 is described the field-trial setup and are presented the results. Their implications are discussed in Section 4, where we also give some of the future work objectives and possibilities. Conclusions are given in Section 5.

2. Fusion networking

The Fusion network is a packet oriented network implementing the principles of integrated hybrid optical networks [11], [12] while employing Ethernet transport over optical wavelengths. The network model is divided into: (1) circuit domain where packets are following a circuit through the network, e.g. a lightpath in the wavelength routed optical network; (2) packet domain where packets are forwarded in a statistical multiplexed manner with a lower priority than on the circuits. However, the two domains are fully integrated and share the same resources, e.g. a common 10Gb/s Ethernet wavelength channel. This differs from the well-known parallel hybrids [10] which use different wavelength channels for separating packet and circuit domains.

The circuit transport in Fusion is named Guaranteed Service Transport (GST) and the packet transport allows Statistical Multiplexing (SM) and is therefore called SM transport. The aim of the GST transport is to provide an Ethernet transport with performance comparable to the Optical Transport Network ITU-T standard (G.709) transport [13]:

1) predictable low and constant latency;
2) no packet loss caused by contention;
3) full transparency also on timing allowing packet-layer synchronization;
4) isolated circuits enabling privacy and performance independence.

The SM transport provides an Ethernet transport comparable to best effort transport in a packet switched network:

1) High network throughput efficiency by employing statistical multiplexing.
2) Applications range from video and voice services to E-mail and Web browsing as the quality of service level depends on dimensioning of the network and the total network load.

2.1 The node description

The schematic diagram of the hybrid node is presented in Fig. 1. The node has two 10 Gb/s Ethernet (10GE) interfaces for the wavelength transport channel. Ten Gigabit Ethernet (GE) interfaces are applied to increase the channel utilization by adding SM traffic. Packets entering a hybrid node are tagged with a Virtual Local Area Network label (VLAN-ID) indicating the type of service. Any SM packet received at the input 10GE interface is dropped to one of the GE interfaces while the GST packets received at the 10GE interface pass through to the other 10GE interface with absolute priority. At the network level, packets classified as GST are forwarded through the network along an end-to-end dedicated path. This traffic bypasses intermediate nodes, requiring only the minimum processing for identifying the packets label as GST stream. Thus, circuit performance properties are provided to this Ethernet traffic class: a low fixed latency and no packet loss as there is no contention in its path. The SM Ethernet traffic is inserted in free time gaps detected between the GST stream of packets. The SM scheduling mechanism in the node ensures that GST packets are left untouched by SM insertion, avoiding PDV and packet loss for GST.

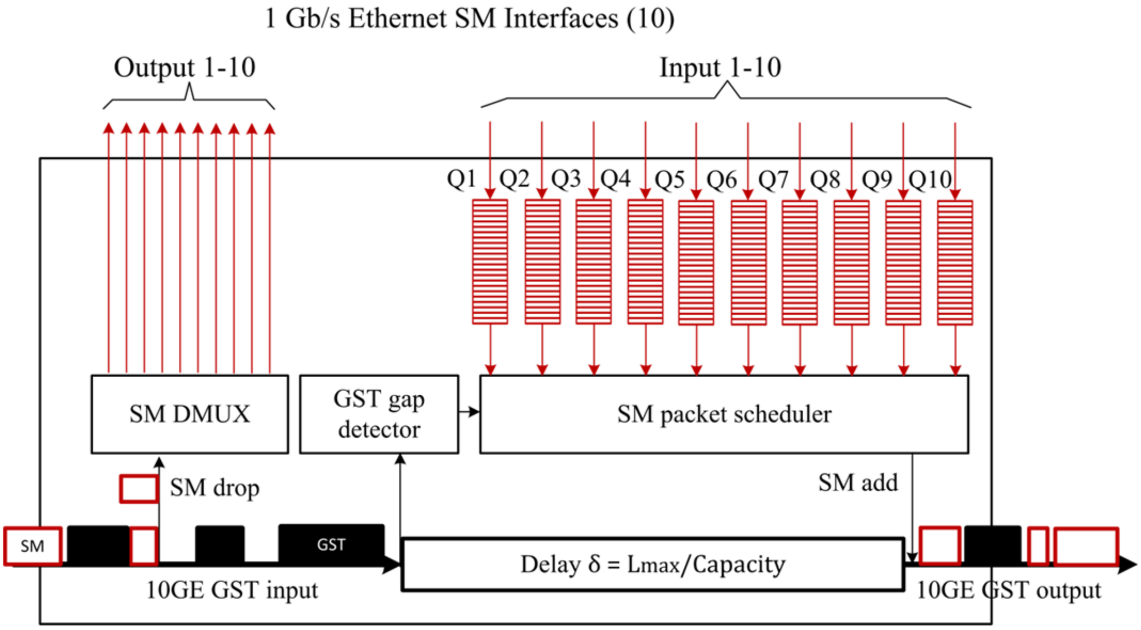

Figure 1 Schematic diagram of the unidirectional (left to right) transport through the fusion hybrid node with ten 1 Gb/s Ethernet interface for local add/drop and one 10Gb/s Ethernet wavelength. The gigabit Ethernet interfaces are depicted logically divided into input and output interfaces to better illustrate the de-multiplexing of SM traffic when entering the node and the scheduling of input traffic in available gaps in the output wavelength.

The only influence the Fusion node has on the GST circuit stream is that bypassing packets are given a fixed delay δ corresponding to the transmission time of a maximum length SM Ethernet packet. For example, in a node with 10 Gb/s channels and with maximum Ethernet frame of 1500 Byte, the delay added at a node is 1.2 µs. Note that the inter-frame timing between the packets in the GST streams in not affected by δ, only the end-to-end delay will increase by $N \cdot \delta$, where N is the number of nodes in the circuits' path. Thus, the delay is deterministic in the network and is applied in the Fusion hybrid node architecture to ensure that:

- Any on-going scheduling of SM packets is allowed to finish without interfering/being pre-empted by GST packets arriving at the input channel. This non-preemptive scheme is energy efficient as processing and transmission of SM packets is done only when the transmission of the packet will be successful.

- The gap-length between GST packets is detected. The SM packet scheduler knowing the free gap, searches in a round-robin manner the head of SM queues e.g. Q1-Q10 in Fig. 1, for a packet smaller than the detected gap. If such packet is found, it is scheduled. If there is leftover space in the gap, the scheduler proceeds until:

1) the gap is full;
2) no suitable SM packet is found in any head-of-queue;
3) all SM queues are empty.

SM packets are therefore forwarded through the network with variable latency, and if the mean load is high due to a high extent of oversubscription, packets will be lost as in a packet switched network.

3. Field-trial setup and results

The objectives of the field-trial, as described in the introduction section, are to demonstrate that the fusion networking provides two Ethernet transport services with these qualities:

1. GST traffic with:
    a. circuit QoS emulating a leased line: low delay and PDV, no loss.
    b. absolute priority and independent of the insertion of additional statistically multiplexed SM traffic;

2. SM traffic that increases the network throughput significantly.

3.1 The field-trial setup

A simple connectivity diagram of the field-trial setup is illustrated in Fig. 2. In the carrier network of UNINETT, the University institutions in Trondheim and Oslo are connected through a wavelength link and two Juniper routers. We inserted two Fusion nodes at each side in between the routers and the wavelength link. One of the 10 Gb/s Ethernet interfaces was connected to the router's interface and the other to the Calient WDM transponder with OTN framing that is used by UNINETT to interconnect the two sites. The Fusion nodes used in our experiments are H1 fusion prototypes developed and provided by TransPacket [14], a Norwegian company developing optical communication systems with the aim of commercializing academic research on the Fusion network.

The production traffic was marked as GST and bypasses through the Fusion nodes. The test-bed has been previously characterized in [15] and a one-way production traffic mirror had been transported. This mirror traffic was doubled to increase the load in the link and at the time of the testing the overall load was quite high because of an important sports championship event in Trondheim. In this work we used the test-bed for carrying the real production traffic between the sites, without interfering on its characteristic and monitoring the performance, i.e. no bit error rates were registered during the duration of the test and no packet losses.

In Fig. 3 is shown the link utilization during a normal working day (without Fusion networking). The statistics are provided by the UNINETT routers. We see that the average daily throughput is approximately 10% of the 10 Gb/s link capacity with a busy hour peak of 2.4 Gb/s in the traffic directed from Trondheim to Oslo. Thus, more than 75% of the link capacity is not used.

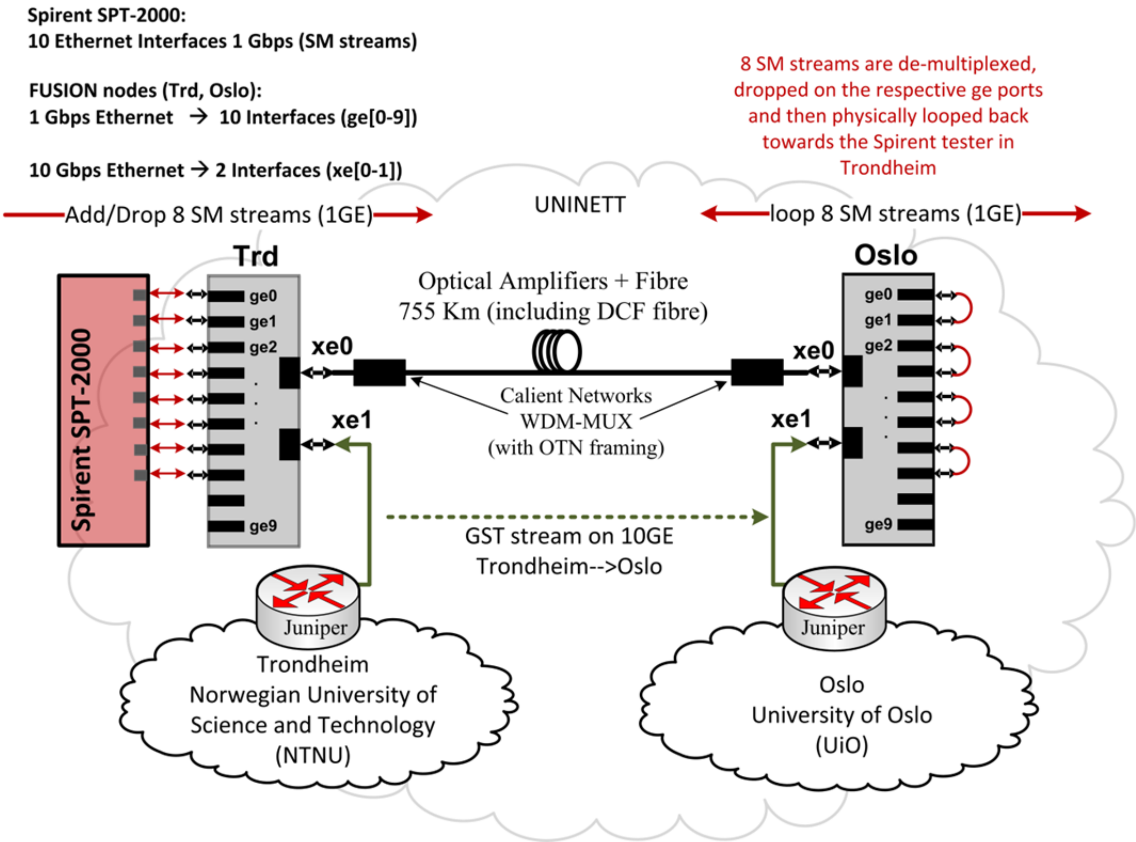

Figure 2. Long-haul field-trial between Trondheim and Oslo through the UNINETT carrier network with two Fusion nodes from TransPacket (http://www.transpacket.com) connecting the sites in NTNU and UiO. The production traffic added by the routers is marked as GST at the xe1 interfaces (10Gb/s) of the fusion nodes and is forwarded transparently and with absolute priority. A Spirent SPT-2000 traffic generator/analyzer adds SM traffic through eight 1Gb/s interfaces. This traffic is physically looped at the node in Oslo to be received at the tester in Trondheim for characterization.

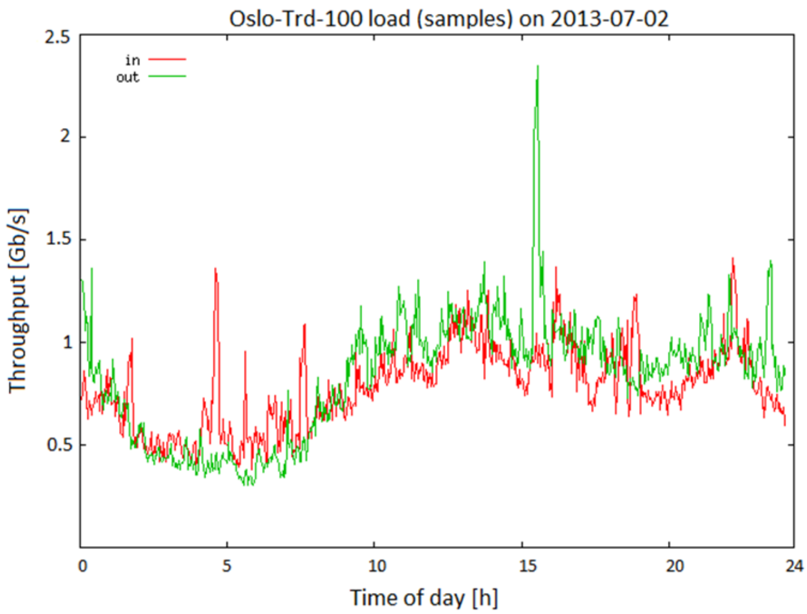

Figure 3. Example of the collected statistics on the Oslo-Trondheim link. The link throughput as a function of time of day.

GST undergoes a fixed delay δ of 8.8 μs at each node in order to avoid preemption of SM packets up to a maximum jumbo packet of length 11000 Byte. At the site in Trondheim, eight SM streams were generated by a Spirent SPT-2000 traffic generator and added to eight GE ports of the fusion node. Each SM stream had an average load of 0.98 on the 1 Gb/s interface (GE) resulting in a total offered SM load of 0.784 on the 10 Gb/s link (10GE). The SM packet length generated by the tester follows an empirical distribution between 40 and 1500 bytes taken from Internet traffic measurements [16]. The SM streams were transmitted through the vacant GST gaps in the channel in the Trondheim-Oslo direction and were looped back to be received and characterized at the traffic analyzer. The SM traffic received at the Oslo node was de-multiplexed to the respective GE ports pertaining to the VLAN tags configuration; then physically looped back to Trondheim and the traffic analyzer. Hence, the SM traffic experienced twice the congestion as a result of the two-way GST traffic. The maximum SM throughput and total channel utilization were measured during 6 hours on the busy/working period of the day.

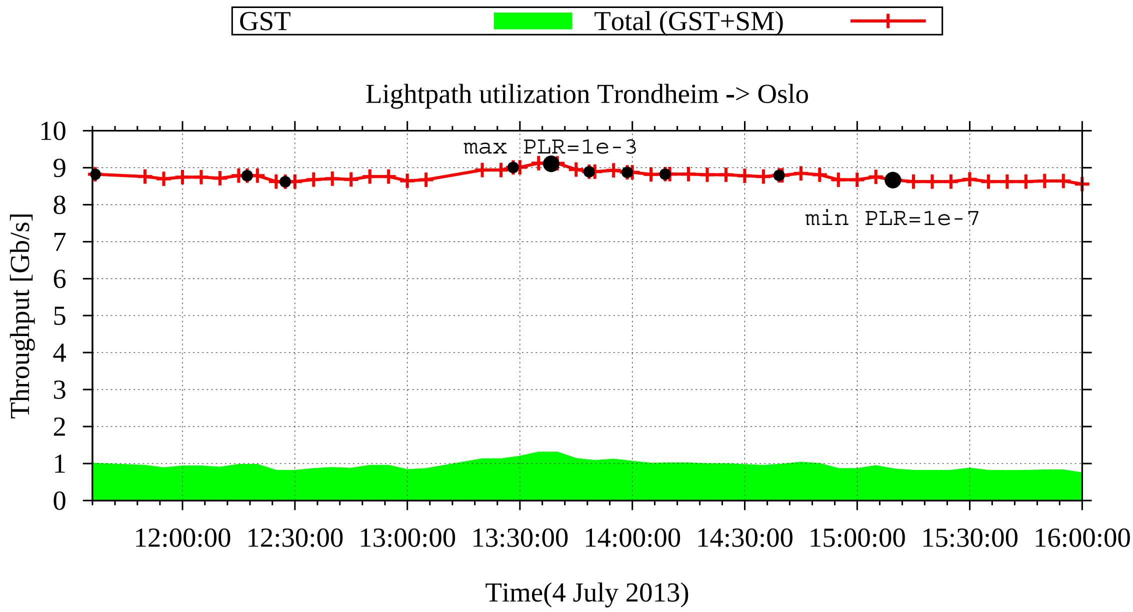

Figure 4. Average link throughput as a function of time of day. The production traffic is carried as GST depicted with the filled lower curve (green). The total throughput with added SM traffic is depicted by the curve (red) to an average of 88% of the capacity. The points are marking the observation of SM losses. GST statistics are collected by the Fusion nodes with 5 min sampling and for SM by the Spirent traffic generator/analyzer with 10 min sampling.

3.2 Results

Before transporting the real production traffic, the fusion network was tested with mirror production traffic for one week, showing no bit errors and no losses on the GST traffic. In addition, a reference characterization of the system was done connecting the Spirent traffic generator to the 10 Gb/s port of the Trondheim node, i.e. the xe1 port connected to the router in Fig. 2. This traffic marked as GST, once received in the Oslo node, was sent back towards Trondheim by physically looping the 10GE ports in the Oslo node. The average end-to-end (one-way) delay was measured to 3792.6 μs, independently of the GST traffic load or the added SM traffic. Knowing that the end-to-end propagation delay on the 755 km fiber of the Trondheim-Oslo link is 3775 μs, results showed that each fusion node added only the fixed delay δ=8.8 μs to the GST traffic. The packet delay variation resolution of the Spirent traffic generator is 10 ns and the added PDV was measured to less than 15 ns. Hence, the Ethernet GST traffic is transported through the Fusion network with circuit quality of service.

In Fig. 4 are presented the field-trial results with real production traffic during a 6 hour time-frame. The maximum achieved throughput is 9.12 Gb/s and is achieved when GST is at its peak. The measured maximum GST rate is 1.28 Gb/s measured at 13:35 and the minimum is 0.72 Gb/s measured at 16:00. The samples where SM has experienced losses are pointed out in circles with a maximum packet loss ratio (PLR) of 1e-3 when GST is at its peak at 13:38 and a minimum 1e-7 of the observed losses. The total throughput averaged through all measured samples is approximately 8.8 Gb/s while the average throughput of the production traffic is 0.93 Gb/s. Hence the link utilization was increased by 78.7% through the deployment of the Fusion nodes. The end-to-end delay of SM streams is dependent not only on the SM load but the congestion from GST traffic as well. In the case that the system is in saturation, when SM streams are experiencing losses, the maximum SM delay is equal to the maximum buffer depth of a GE interface. In the H1 nodes this buffering capacity is 268 Mb, corresponding to a maximum buffering delay of 268 ms. The SM delay results at the points where losses are observed, see Fig. 4, are much lower with an average RTT of 17 ms.

4. Discussion and Future Work

However diverse and wide the QoS requirements of network services and applications might be, the fundamental needs are performance related: bandwidth, delay, packet delay variation and packet loss rate. The Fusion architecture enables a low-cost standard Ethernet network with differentiated classes of service by a combination of these performance parameters.

The experiment results show that Ethernet-based Fusion networking can give a circuit service with hard QoS which confirms the suitability for deployment in the mobile back-haul network. The GST class is suitable to transport time-sensitive information. It can carry the synchronization information through adaptive clock recovery schemes e.g. the PTP protocol, because the end-to-end delay was proved to be deterministic and the only added delay we observed was the applied fixed delay δ per node. Furthermore, the delay δ was set for allowing the transmission of an SM jumbo frame of 11000B but it is possible to optimize it according to the SM traffic distribution that the network engineer expects in its network and the delay budget of the services.

In the same back-haul network, the SM service class can then be used to transport the not time-critical data traffic with the ever-increasing bandwidth requirements in the mobile network [18]. Within the SM class, different QoS differentiation mechanisms might be implemented. In [12] the authors propose the division of SM traffic into real-time and best-effort classes. This scheme is suitable for transporting strict real-time traffic from e.g. multimedia applications like VoIP, that are affected by the playback latency effect; and best effort web-traffic, where latency may not be an issue, e.g. web page download from Internet requires only the complete file (low packet loss) with non-critical delay. In addition, high quality video communications such as 3D-TV, digital video broadcasting and tele-surgery are very demanding [19-21] and fit well with the GST transport.

Results from this field-trial and the ones presented in [15] and [17] show that the SM performance, i.e. maximum achievable SM throughput and average delay, is dependent on the GST load in the links. Reactive load-balancing mechanisms should be applied in order to maintain the SLA requirements in the SM class. Besides, the GST traffic load and pattern is related to the time of day and large-amount data transfers through the SM class can be scheduled for less loaded hours. For example, the additional SM capacity can be used in data-center networks for scheduled data back-up.

The Fusion node is able to aggregate GST traffic from one of the 1 Gb/s port to the 10 Gb/s port with bounded aggregation delay [17]. The possibility of extending this functionality in the future to multiple 1-to10 Gb/s ports would allow provisioning Ethernet circuit services with finer granularity.

5. Conclusions

In this work we presented a long-haul Ethernet-based Fusion network field-trial with real production traffic in the carrier network of UNINETT. The results confirm that the GST transport provides to Ethernet a transport service with circuit QoS: zero packet loss, low constant delay much lower than the propagation delay in the fiber, and low PDV in the range of nanoseconds. Furthermore it has a low processing overhead and enables efficient transparent router bypass, transport of time-sensitive traffic and packet layer synchronization information. Its performance is independent of the insertion of additional statistically multiplexed traffic in the same optical medium.

High throughput efficiency was demonstrated by adding Ethernet packet-switched SM traffic. Results show that the capacity utilization was increased by 78.8% to a total average link utilization of 88% with a peak of 91%.

The field-trial results confirmed that the Fusion technology is able to deploy Ethernet in carrier networks, maintaining its important low-cost high-throughput feature while enabling the new service of circuit transport on the same physical resources.

References

[1] A. Meddeb, "Why Ethernet WAN transport?," *IEEE Communications Magazine,* Vol. 43(11), pp. 136-141, 2005.

[2] K. Fouli, M. Maier, "The road to carrier-grade Ethernet," *IEEE Communications Magazine,* Vol. 47(3), pp. 30-38, March 2009.

[3] F. Luyuan, R. Zhang, M. Taylor, "The evolution of carrier ethernet services-requirements and deployment case studies [next-generation carrier ethernet]," *IEEE Communications Magazine,* Vol. 46(3), pp.69-76, March 2008.

[4] P. Briggs, R. Chundury, J. Olsson, ‘Carrier Ethernet for mobile backhaul’, *IEEE Communications Magazine*, Vol. 48(10), pp. 94-100, 2010.

[5] Z. Ghebretensaé, J. Harmatos, K. Gustafsson, "Mobile broadband backhaul network migration from TDM to carrier Ethernet," *IEEE Communications Magazine,* Vol. 48(10), pp.102-109, October 2010.

[6] L. Cosart, “NGN packet network synchronization measurement and analysis," *IEEE Communications Magazine,* Vol. 49(2), pp.148-154, 2011.

[7] S. Bidkar, S. Mehta, D. Bhamare, N. Bajaj, A. Medhekar, A. Gumaste, "Circuit performance in a packet network: Demonstrating integrated Carrier Ethernet

Switch Router (CESR) + Optical Transport Network (OTN)," *International Conference on Computing, Networking and Communications (ICNC),* pp.400-407, 28-31 Jan. 2013.

[8] C. Estevez, S. Angulo, A. Abujatum, G. Ellinas, Liu Cheng, G. Chang, "A Carrier-Ethernet oriented transport protocol with a novel congestion control and QoS integration: Analytical, simulated and experimental validation," *IEEE International Conference on Communications (ICC),* pp.2673,2678, 10-15 June 2012.

[9] The Metro Ethernet Forum, "Ethernet Services Definitions - Phase II", MEF 6.1, April 2008.

[10] C.M. Gauger, J.P. Kuhn, E.V. Breusegem, M. Pickavet, P. Demeester, "Hybrid optical network architectures: bringing packets and circuits together," IEEE Comm. Magazine, Vol 44(8), pp. 36-42 (2006).

[11] S. Bjørnstad, D.R. Hjelme, N. Stol, "A Packet-Switched Hybrid Optical Network with Service Guarantees," *IEEE Journal on Selected Areas in Communications*, Vol. 24(8), 2006.

[12] N. Stol, C. Raffaelli, M. Savi, "3-Level integrated hybrid optical network (3lihon) to meet future qos requirements," in *Proc. Of IEEE GLOBECOM*, Houston, Texas, USA, Dec. 2011.

[13] ITU-T Recommendation G.709/Y.1331 "Interfaces for the Optical Transport Network (OTN)," February, 2012.

[14] "TRANSPACKET H1; a fusion networking add-drop muxponder," [Online]: http://www.transpacket.com/fusion-h1/fusion-ethernet-h1/, last checked July 2013.

[15] S. Bjornstad, R. Veisllari, and K. Bozorgebrahimi, "Long-haul Ethernet-Transport Integrated Packet/Circuit Hybrid Network Field-Trial," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference (OFC/NFOEC)*, paper PDP5A.8, 2013

[16] CAIDA, Traffic Analysis Research, [Online]: http://www.caida.org/research/traffic-analysis/pkt_size_distribution/graphs.xml, last checked July 2013.

[17] R. Veisllari, S. Bjornstad, and K. Bozorgebrahimi, "Integrated Packet/Circuit Hybrid Network Field-Trial," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference (OFC/NFOEC)*, paper NW4I.5, 2013.

[18] Cisco Whitepaper, "Cisco Visual Networking Index: Forecast and Methodology, 2011-2016", [Online]. Available: http://www.cisco.com/, last checked July 2013.

[18] General Services Administration, "Networx Service Level Agreement (SLA) Management Guide Version 2.0," April 2009.

[19] 411- ITU-T FG IPTV "IPTV QoS/QoE Metrics," January 2007.

[20] ETSI EN 300 744, "Digital video broadcasting (DVB): frame structure, channel coding and modulation for digital terrestrial television (DVB-T)," 2004.

[21] ITU-T recommendation Y.1541 (12/11), "Internet protocol aspects Quality of service and network performance", December 2011.