High-Speed Transport and Aggregation for Ethernet Fronthaul with Low and Bounded Delay

Raimena Veisllari[1], Steinar Bjornstad[1,2], Mickael Fontaine[1], Carla Raffaelli[3]
[1]TransPacket, Hoffsveien 21-23,0275 Oslo, Norway
[2]IIK, Norwegian University of Science and Technology, Trondheim/Simula Oslo MET, Norway
[3]DEI, UniBO, Bologna, Italy
Email:{raimena}@transpacket.com

Abstract: 3-node integrated packet/circuit Ethernet network experiment demonstrates aggregation and add/drop of 10Gb/s fronthaul links on a 100Gb/s path. Bounded 5.9µs maximum end-to-end delay and 1.24µs PDV is achieved, even when combining with less delay-sensitive traffic.
OCIS codes: (060.4253) Networks, circuit-switched; (060.4259) Networks, packet-switched

1. Introduction

The ability to centralize and virtualize the radio access network (c/vRAN) architecture is one of the key enablers for cost-efficient and scalable 5G deployment. The different functional splits and function placement of the mobile protocol stack between the remote radio unit (RU), distributed unit (DU) and the centralized unit (CU) create a new mobile transport network [1], with fronthaul (FH) defined between RU and DU, midhaul (MH) defined between DU and CU, and backhaul (BH) between CU and the core network. Each of these *x*-hauls poses diverse, sometimes conflicting, quality of service requirements to the underlying transport network. For lowering deployment costs, adapting, and scaling the transport network, it is desirable for the service provider to be able to serve all front-, mid- and back-hauls in an integrated infrastructure, supporting a highly flexible vRAN architecture, where virtualized functions can be deployed around the different mobile edge sites.

In this context, research and standardization work has focused on Ethernet, whose high-bitrate interfaces and transceivers are available at low cost and continuously enhanced and standardized. Ethernet packet-based technology efficiently matches with the packet-based user data traffic and variable bit rates of mobile traffic by applying statistical multiplexing. Being already widely deployed in the backhaul segment, Ethernet enables interoperability with the current 4G deployments. As a consequence, standardization effort has been focusing on packetization of radio/mobile data over Ethernet, e.g. eCPRI [2] and IEEE 1914.3 Radio over Ethernet [3], and transport over Ethernet networks, e.g. IEEE TSN 802.1CM [4] and IEEE 1914.1 [5]. However, the challenge with Ethernet, as with packet-switching, is that the gain in resource utilization comes with the cost of statistical performance: packet delay and packet delay variation (PDV) are dependent on statistically multiplexed service traffic patterns, and need to be accounted for. This is a direct issue for fronthaul, having the strictest maximum one-way delay, i.e. including peak-to-peak PDV, of 100µs [2]. Early deployments comply to the delay requirements by using isolated point-to-point connections, e.g. dedicated fibers or wavelengths for each fronthaul stream, where the speed of light in the fiber gives the transport delay. However, these point-to-point link topologies might not be viable, scalable or cost-effective in all cases, e.g. fiber and wavelengths might be a scarce and/or expensive resource. Furthermore, the need for more capacity and the use of millimeter wave links (and their smaller coverage) will drive densification of RU/DUs and small cells. Thus, aggregation and switching mechanisms with low and bounded delay become important for deploying scalable and cost-effective Ethernet mobile transport networks.

In this work we present and experimentally evaluate mechanisms for low and bounded delay Ethernet aggregation, transport, and add/drop based on integrated hybrid (packet/circuit) optical networks (IHON) [6], also known as FUSION. IHON enables a multi-service path supporting: (1) guaranteed service transport (GST) service class with fixed delay and ultra-low packet delay variation (PDV) independent of load, and (2) high throughput efficiency through a lower-priority statistically multiplexed (SM) service class. In previous work [7] the suitability of the GST class for carrying fronthaul in a 100G Ethernet path through multiple nodes has been demonstrated, while sharing the path with mid/backhaul streams served by the SM class. Furthermore, in [8], a round robin bounded delay scheme was proposed for asynchronous aggregation of multiple, equal priority fronthaul streams into a 100G GST class path. In this work, the bounded delay scheme is further extended, demonstrating aggregation and add/drop of multiple FH streams with different bandwidth requirements, through a multi-node 100G path with low and bounded delay. The mechanisms are evaluated through an experimental IHON demonstration testbed, emulating a 3-node network. The 100G path is shared with mid/backhaul streams, demonstrating the feasibility of a fully integrated Ethernet mobile transport network fulfilling the diverse latency bounds of each *x*-haul service.

2. Mechanisms for Deterministic Aggregation with Low and Bounded Delay

The main IHON mechanism extended for bounded delay aggregation and add/drop of FH traffic, is the GST bypass – SM insert [6], illustrated in Fig.1(b). The high priority stream at the transport path's 100GE input (GST) is bypassed

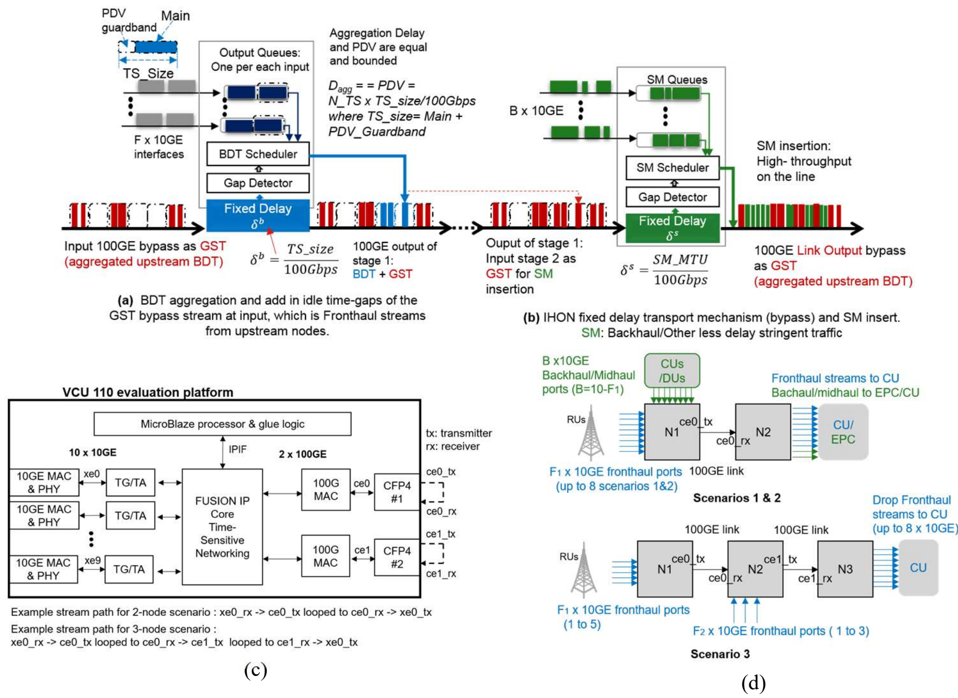


Table 1. Parameters

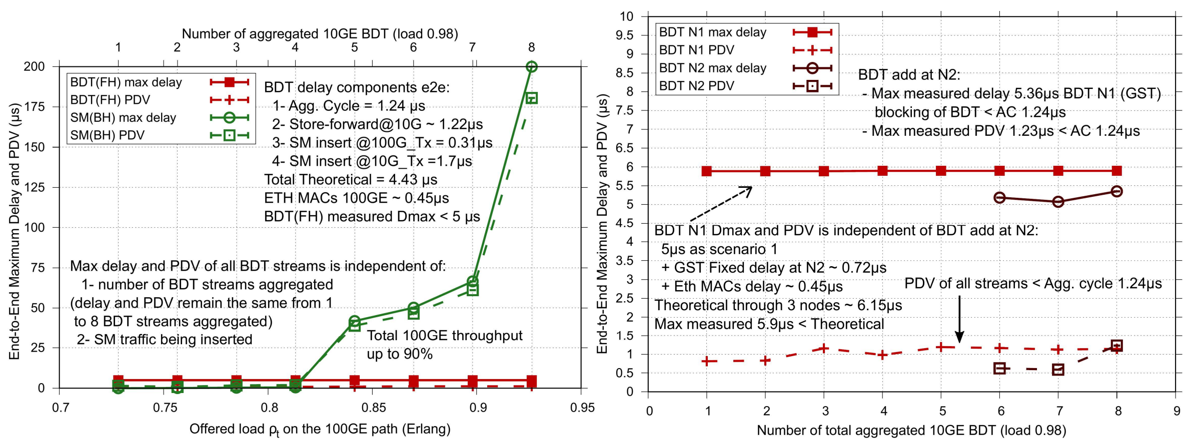
Parameters	Desc./Value
BDT fixed delay	$\delta^b = \frac{TS\ main}{100Gb/s}$ ~0.25µs
SM fixed delay	$\delta^s = \frac{1550B}{100Gb/s}$ ~0.2µs
TS main	1588B
PDV guardband	300B (24ns@100GE)
TS size	1588B+300B
Aggregation Cycle AC	$\frac{8\ x\ TS}{100Gb/s}$ ~1.24µs
FH 10GE ports at node n	F_n
BH/MH 10GE ports	$B = 10 - F_1$
FH load at 10GE port	ρ_x^f=0.98
BH/MH load at 10GE port	ρ_x^b=0.7
Total load 100GE path	ρ_t

Fig.1: (a) BDT aggregation in the IHON node with three classes of services: GST, BDT, and SM. (b) SM insertion in idle time gaps of the bypassing stream at the 100G output. (c) Experimental setup, 100GE ports are looped back for (d) emulating two and three node networks; scenario 1 & 2: BDT aggregation w/o and w/ SM (10GE streams looped back at port ce0 and dropped at the respective 10GE ports); scenario 3: BDT aggregation at N1 (looped back at ce0 sent to ce1) and BDT add at N2 (looped at ce1 and all dropped at 10GE ports).

to the 100GE output with a fixed delay δ, corresponding to the service time of a maximum transmission unit (MTU) SM packet to be inserted. This creates a time-window before the GST packet reaches the output, enabling a look-ahead for a gap detector to identify idle GST inter-packet gaps. Lower priority SM packets are scheduled using this gap information and transmitted only if they fit the gap, i.e. without affecting the timing of the packets in the bypassing GST stream. In the xRAN context, the GST class fits with FH, and the SM class with MH/BH. This architecture is further extended for aggregation and add/drop of multiple lower bitrate FH streams in the transport path. Currently, the bypass-insert mechanism is implemented in cascade as illustrated in Fig. 1(a,b), enabling a third service class: bounded delay transport (BDT), with higher priority than SM. Once a suitable gap is found, the BDT aggregator serves the correct BDT input queue; once the BDT streams are served on the 100GE transport path, they are classified as GST on the ingress of the downstream nodes. The transport delay is thus low and fixed to the number of nodes traversed times the configured fixed bypass delays. E.g. the GST stream is bypassed first through the fixed delay δ^bapplied for BDT insertion, and then both the inserted BDT and bypass GST go through a second fixed delay δ^s applied for SM insertion. The bounded delay transport (BDT) aggregation and add is illustrated in Fig.1(a). Each BDT stream is allocated one or more equal size virtual time-slot(s) (TS), depending on the peak bandwidth requirement. The TS contains a *main* part, the maximum allowed inserted burst size, and a configurable PDV guard-band, left unused. An aggregation cycle (AC) time is then defined in the network nodes based on: (1) TS size, relating to the smallest required peak bandwidth of the streams, and (2) total number of TS - each BDT stream must be allocated an integer number of TS(s) based on the TS bandwidth granularity. The maximum aggregation delay experienced by a BDT stream is thus bounded to the AC time: if the packet just missed its TS, it will be served on the next cycle. This relates also to the maximum PDV experienced at the aggregation node. The BDT aggregation/add relies on finding an idle gap for the payload to be inserted in a time-slot and therefore is different from other slotted schemes. In an asynchronous network, the configurable PDV guard-band ensures that any delay variation induced by imperfect physical layer implementation, e.g. Ethernet MAC, of the traversed nodes does not shrink the time-slot “allocated” to the downstream nodes in a way that sufficient unfragmented bandwidth remains to add BDT.

3. Experimental Results and Discussion

The IHON node used in the experimental setup, illustrated in Fig. 1(c), is implemented on Xilinx VCU110 Ultrascale FPGA evaluation board, including Ethernet PHY/MACs for the respective 10/100GE ports, and the IHON mechanisms implemented by TransPacket. As one node is used for emulation, no clock-offset is present in the demonstration. The PDV guard-band is therefore used for protecting against the physical layer PDV. Each 100GE port has its own scheduling mechanism on transmit side, thus demonstrating asynchronous gap-filling BDT. For emulating all traffic streams, ingress traffic generators and egress traffic analyzers are implemented on the 10G ports in the FPGA. The MTU for all streams is fixed to 1522B: 1518B is the Ethernet maximum sized packet and 4B the VLAN tag used for differentiating traffic type. The loads are fixed by fixing the inter-packet gaps, see Table 1. The nodes are configured with AC=1.24µs for eight 10GE FHs in the 100GE path, main=1588B and PDV guard-band 300B (24ns). Three test scenarios are considered: (1) only BDT aggregation from 1 to 8 10GE streams on N1 and dropped on N2 - verifying that BDT streams are isolated from each-other; (2) as 1 with SM insertion where number of SM streams is (10-FH) – verifying that BDT streams are isolated from SM and a high throughput is achieved on the 100GE path; (3) BDT aggregation from 1 to 5 10GE streams at N1, and BDT add of 1 to 3 10GE streams at N2, all terminated/dropped at N3 – verifying that aggregated BDT is isolated from BDT add at downstream nodes. Maximum end-to-end (e2e) delay (*Dmax*) and PDV for the first two cases are illustrated in Fig. 2(a) while scenario 3 in Fig. 2(b). BDT delay is demonstrated to be bounded to the aggregation cycle 1.24µs + fixed processing delays in the IP core + transmission delay + Ethernet MACs measured reference delays, for all three cases. The maximum measured PDV of 1.23µs is consistent with the theoretical bound equal to the aggregation cycle of less than 1.24µs for both BDT aggregation and add/drop, independent of number of BDT streams aggregated/added and SM streams inserted. Note that for scenario 3, *Dmax* for F_1 streams increases according to the bypass delays experienced at N2, less than 1µs, where 0.72µs is the fixed bypass and processing delays, and delay added by the 100GE MACs (measured reference delay up to 0.45µs). Hence, for a 2-node network the BDT e2e maximum delay added by the nodes is less than 5µs, where 1.24µs is the AC, while for a 3-node network less than 5.9µs, where 0.71µs is the GST bypass delay at N2 and less than 0.45 µs is the additional 100GE MAC delay. The total maximum throughput is found by inserting a worst-case SM traffic pattern of fixed MTU packets - smaller packets distribution would fill better the gaps [5]. A 100GE path throughput up to 90Gbps was achieved, where the SM delay is well within the range of 500µs [1], showing that results are not only fitting for BH but also for the more demanding MH traffic.



(a) (b)

Fig.2: (a) Maximum (of all BDT streams) delay and PDV as a function of the total offered load in the 100G path, with FH as on/off sources with offered load 0.98 on the 10G ports emulating CPRI opt.7 over Ethernet; measured BDT performance is equal for both 2-node scenarios 1 w/o SM and 2 w/ SM at 0.7 load on (10-F_1) ports. (b) BDT max delay and PDV for 3-node scenario 3.

4. Conclusion

IHON bounded delay aggregation and add/drop of 10 Gb/s streams into 100 Gb/s Ethernet, suitable for transport of Ethernet fronthaul streams coming from multiple radio units, was experimentally demonstrated for the first time. The FH path was bridged through three emulated nodes, and multiplexed with less delay-sensitive traffic, e.g. BH/access services. Zero packet loss, maximum delay bounded to 5.9 µs and peak PDV less than 1.24µs – independently of traffic load - were demonstrated, leaving 94 µs of a 100 µs fronthaul delay budget to transmission delay.

5. Acknowledgement

This work is partially funded by research project 5G-PICTURE supported by the European Horizon 2020 initiative.

6. References

[1] 3GPP Tdoc R3-161380, "Motivation for standard interface between central and distributed units".
[2] CPRI Consortium, "eCPRI Specification V1.0," Aug. 2017.
[3] IEEE Std. 1914.3, "Radio Over Ethernet Encapsulations and Mappings," Sept. 2018.
[4] IEEE Std. P802.1CM, "Time Sensitive Networking for fronthaul".
[5] IEEE Std. P1914.1, "Standard for Packet-based Fronthaul Transport Networks," Sept. 2018.
[6] S. Bjornstad et al., OFC 2013, post deadline paper PDP5A.8.
[7] R. Veisllari et al., ECOC 2017, paper M2A2.
[8] R. Veisllari et al., OFC 2018, paper M4I2.