Minimizing Delay and Packet Delay Variation in Switched 5G Transport Networks

S. Bjørnstad, R. Veisllari, D. Chen, F. Tonini, C. Raffaelli

Abstract— **A key requirement in 5G networks is the demand of minimizing transport delay. In this paper we propose and explore novel mechanisms for minimizing packet delay and delay variation in 5G mobile transport networks. Differently from mechanisms described in the literature and standardization for deterministic delay aggregation, a novel method using a time-window combined with a timeout-based approach, is proposed. The goal is to achieve bounded delay aggregation of traffic, suitable for application in fronthaul transport. We propose an asynchronous approach to deterministic networking, enabling Bounded Delay Transmission (BDT). Exploration of parameters for controlling the maximum delay and Packet Delay Variation (PDV) of the aggregated traffic is performed by simulation, and practical demonstration is shown in an experiment. Using the proposed mechanisms, we show that the delay and PDV can be bounded below 20 microseconds with up to 85-90% carried load of the aggregated flows. Hence, we demonstrate their suitability for delay sensitive fronthaul transport and future applications in 5G networking. Experimental results demonstrate that the practical implementation of the mechanism is viable. Thus, it motivates the extensive analysis performed through simulations.**

Index Terms—**5G, delay, eCPRI, fronthaul, IHON, TSN, PDV**

I. INTRODUCTION

Mobile fronthaul is a key enabler for the 5G transport network in 5G mobile networks. By centralizing the higher layers of baseband processing functions, currently residing in the radio, a more cost-efficient network design

S. Bjørnstad is with Tampnet AS, Stavanger, Norway, IIK NTNU Trondheim, Norway, Simula Oslo MET, Norway. steinar.bjornstad@item.ntnu.no
R. Veisllari is with TransPacket AS, Oslo, Norway. Raimena.Veisllari@transpacket.com
D. Chen is with Nokia, Naperville, Illinois, USA. David.1.Chen@nokia.com
F. Tonini and C. Raffaelli are with the Alma Mater Studiorum, Universita' di Bologna, Italy. f.tonini@unibo.it, carla.raffaelli@unibo.it

can be accomplished [1, 2]. The functions are then split into a Base Band Unit (BBU) and a Remote Radio Head (RRH), which are separated and connected by a low delay/delay variation and high throughput network segment called fronthaul. Until recently, Common Public Radio Interface (CPRI) and Open Base Station Architecture Initiative (OBSAI) have been the preferred protocol formats for fronthaul. However, considerable effort is being devoted to migration to Ethernet transport for fronthaul deployment. In fact, while CPRI and OBSAI are protocols designed for mobile fronthaul only, Ethernet is a packet-switched standard in constant enhancement, widely used in both telecom and enterprise networks.

Recent initiatives for introducing Ethernet in fronthaul transport include studies in the CPRI Cooperation with the new industry standard enhanced CPRI (eCPRI) [3], IEEE 1914 working group with IEEE 1914.1 Next Generation Fronthaul Interface [4], IEEE 1914.3 Radio over Ethernet [5], xRAN Fronthaul Specification [6], and IEEE 802.1CM Time Sensitive Networking (TSN) for fronthaul [7]. However, applying Ethernet in mobile fronthaul puts a new level of performance requirements, especially for delay, delay variation, packet loss, and reliability parameters [8]. For example, in IEEE 802.1CM, the maximum end-to-end one-way delay for fronthaul is set at 100 µs, including fiber and Ethernet bridge delay. Delay requirements are therefore identified to be one of the most challenging parameters to fulfill using packet-switched Ethernet systems since they were not originally designed for delay-sensitive networks.

In this paper, novel mechanisms for bounding the delay and the Packet Delay Variation (PDV) in packet-switched networks are proposed and analyzed. Differently from our earlier works on aggregation using containers and synchronization packets for achieving fixed delay transport [9], the mechanism proposed here, namely Time-Window with Timeout (TWT), is asynchronous and can achieve a Bounded Delay Transmission (BDT) that does not rely on synchronization packets. Additionally, it is tolerant for PDV added by nodes in the network path. In [10] we have proposed aggregation and add/drop of traffic streams using a Time-Window (TW) approach and performed a first experimental verification. In this paper we extend this work by simulating both the TW and the new TWT mechanism within an extensive parameter space and compare baseline simulation results with experimental results.

The outline of the paper is as follows: In section II we

review emerging standards and already proposed mechanisms to control delay. In section III we discuss the delay components in packet switched fronthaul and how these components influence PDV and delay. In section IV the TW and the new TWT bounded delay mechanisms are described. In section V numerical evaluations obtained by simulation are presented and discussed. In section VI experimental results are shown demonstrating baseline performance for the TW mechanism and motivating the exploration of the TWT mechanism. In section VII conclusions of the work are drawn.

II. EARLIER WORKS ON MECHANISMS FOR CONTROLLING DELAY IN PACKET-SWITCHED FRONTHAUL

A. Earlier works on fixed delay mechanisms

Mechanisms for enabling deterministic delay can be found both in research literature and in standardization. In literature, solutions related to the jitter control while encapsulating CPRI flows into Ethernet frames have been proposed [8]. Furthermore, several approaches to control and minimize delay by applying synchronization in packet switched networks have been described in the literature. In the Time Shared Optical Network (TSON) [11], contention-free optical switching and transport of time-slots across one or multiple wavelengths per service is proposed. The ability of TSON to transport several different protocol formats, like Ethernet and CPRI has been demonstrated [12]. In [13] "optical Ethernet (OE)" a method using containers with Forward Error Correction (FEC) in time-slots is proposed. A key property of OE is that intermediate nodes only examine the container header for distinguishing between transit traffic and traffic destined for the local node, while transit containers are forwarded in a cut-through-like manner without adding queuing delay.

Different from the time-slotted approaches mentioned above, Integrated Hybrid (packet/circuit) Optical Networks (IHON) enables a fixed delay traffic class combined with a statistically multiplexed traffic class. This has been proposed and experimentally demonstrated in [9]. There are two mechanisms: a priority mechanism, and an aggregation mechanism requiring synchronization. The priority mechanism enables a high priority traffic stream to be processed with fixed delay while receiving absolute priority over lower priority streams.

Fig.1 illustrates the priority mechanism, and shows how packets of the lower priority Statistically Multiplexed (SM) traffic are inserted in gaps, i.e. Time-Windows (TWs), between the packets in the high priority Guaranteed Service Transport (GST) traffic stream. SM packets are inserted only when the TW is sufficiently large to fit the packet. The SM packet selector searches through the first packet in each of the N queues for a packet of suitable size for any available TW in the GST stream. For observing that the TW is sufficiently large for SM packet insertion, the GST traffic receives a fixed delay corresponding to the duration of a maximum sized SM packet. For Ethernet networks this will correspond to the duration of a Maximum Transmission Unit (MTU) packet. Since all GST packets receive the same

fixed delay, and SM packets are inserted only in fitting TWs, the number of bytes between the GST packets can be preserved, inducing no PDV on the GST traffic. This enables full isolation from the SM traffic on the GST traffic. The mechanism relies on monitoring packets and gaps arriving in the GST stream and does therefore not rely on any synchronization of the network. In the illustration, the smaller SM packet in the first queue arrives after the larger one in the second queue, but is transmitted earlier because a small gap first becomes available. Experiments implementing IHON on Ethernet do however show some PDV being induced by the Ethernet MAC [9]. However, the PDV typically corresponds only to a fraction of the duration of an MTU packet.
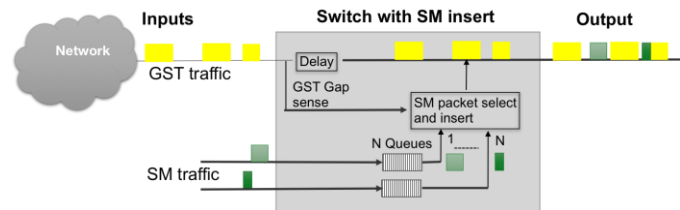

Fig. 1. Asynchronous mechanism for insertion of lower priority Statistically Multiplexed (SM) traffic in a high priority time-sensitive Guaranteed Service Transport (GST) traffic stream.

Deterministic end-to-end delay through a network can be achieved by combining the described priority mechanism with the IHON aggregation and de-aggregation mechanism. The aggregation mechanism is based on aggregating packets in virtual containers starting with a synchronization and control information packet. When aggregating a packet stream in a container, the stream is sliced in pieces fitting into fixed length containers. Both packets and gaps of the stream are preserved and fitted into the container, enabling an exact reconstruction of the streams at the de-aggregation side, without adding PDV. The containers from the streams being aggregated are then multiplexed container by container into a higher bitrate stream. Hence, nodes involved in aggregation and de-aggregation rely on being synchronized. An experimental demonstration using Ethernet, implemented in an FPGA, has proven the viability of the approach [9].

B. Standardization on controlled-delay Ethernet

In standardization, especially within Ethernet, there has been a high activity on mechanisms for controlling and minimizing delay. The IEEE 802.1 working group standardizes Ethernet and mechanisms for Time Sensitive Networking (TSN). A comparison of TSN and IHON mechanisms is found in [14]. Both TSN and IHON contain different mechanisms for both priority and aggregation ensuring deterministic QoS. The TSN priority mechanism is based on preemption and does not rely on synchronization in the network. The IEEE 802.1Qbu [15] defines how lower priority Ethernet frame transmission (preemptable traffic) can be interrupted by high priority (express) traffic frames. The principle is illustrated in Fig. 2.
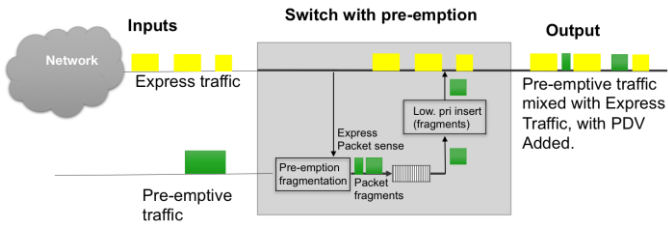
Fig. 2. Preemption mechanism as defined in IEEE 802.1Qbu. An additional delay is added to the high priority express traffic while waiting for the lower priority traffic to be preempted.

Since the lower priority frames can be fragmented, the transmission of a low priority frame can continue after the interruption from a high priority frame. While the mechanism minimizes delay for the express traffic, some PDV is added. Preemption is performed only if 60 bytes or more of the preemptable frame have been transmitted and 64 bytes or more, including the frame's 4 Byte CRC, remain to be transmitted. Including 7 Byte Ethernet preamble, 1 Byte start frame delimiter, and 12 Byte minimum inter-frame gap, 155 (60+63+7+1+12+12) Bytes (1240 bit) of delay may be added [7]. Since the delay is zero if no preemptable packet is present, the PDV of the high priority stream corresponds to the duration of 155 Bytes. Furthermore, as MAC addresses are not added to any packet fragments, the mechanism only works hop-by-hop between bridges supporting the 802.1Qbu standard.

The IEEE 802.1Qbv [16] enhancement for scheduled traffic defines a method based on time-aware scheduling to manage link access by a set of queues associated to input links. For each queue, there is a transmission gate that allows frames to be transmitted when the gate is open and blocks frames from transmission when the gate is closed. By going round-robin, opening and closing the gates, a deterministic aggregation of packets from the queues into a higher bitrate link is enabled. This allows for a bounded delay approach, but differently from the IHON aggregation mechanism, it does not preserve the gaps between the packets in the streams while aggregating them. Hence, if packet streams are arriving asynchronously, a fixed delay as in IHON cannot be achieved. Furthermore, IEEE 802.1Qbv requires synchronization. A global network synchronization based on the IEEE 1588 [17], Precision Time Protocol (PTP), is proposed.

Another mechanism defined by IEEE, not requiring synchronization while still bounding delay, is the IEEE 802.1Qcr Asynchronous Traffic Shaper (ATS) [18]. The shaper assigns an eligibility time to all incoming packets. These eligibility times are applied for selecting packets for scheduling. Packets are dropped if they stay too long within the bridge, i.e. beyond a predefined residence time. This mechanism has the ability to control and reduce the maximum delay through the eligibility time assignment. Furthermore, the eligibility time assignment may be performed according to the QoS marking of the incoming packets, enabling QoS differentiation on delay. Assigning an eligibility time implies a set of calculations to be performed for each packet. The mechanisms proposed in this paper are for asynchronous networks. As for the Qcr mechanism they aim at reducing maximum delay. A key difference from the IEEE 802.1Qcr mechanism is that per packet computations are not needed, easing implementation of the mechanism for processing of high bitrate streams like 100 Gb/s and above.

III. Delay Elements in Ethernet-based Fronthaul Networks

When an Ethernet switch aggregates packet traffic, statistical multiplexing is applied resulting in a set of delay components being added. Assuming fronthaul traffic being generated according to the CPRI standard, mapping is needed to encapsulate CPRI frames into Ethernet frames [8]. This introduces a fixed serialization delay contribution that is a function of the CPRI source traffic speed and the Ethernet frame length. The Ethernet node introduces a further processing delay depending on the input interface speed, if store-and-forward operation is applied. A further contribution is represented by the transmission delay, i.e. the time needed to transmit the Ethernet frame, including its header, on the output interface: this is given by the ratio between the length of the frame (bits) and the speed (bit/s) on the output interface. If traffic bypasses through interfaces having the same speed, cut-through operation is shown to reduce the delay down to a few ns, e.g. 6.4 ns as shown in [8]. In addition to these contributions, which are constant, a variable contribution is represented by queuing delay, originating from statistical multiplexing and contention. Hence an Ethernet frame carrying CPRI traffic is affected by both the constant delay and variable delay components added at each node. This delay is accumulated node-by-node along a multi-hop path. Furthermore, the propagation delay due to the propagation time in the fiber links between each node pair must be taken into account. This added delay corresponds to 5 microseconds per km of fiber.

The maximum end-to-end delay must be compliant with the service constraints. Those services that are not tolerant to delay variation, e.g. CPRI, require compensation based on using a playout buffer. The playout buffer, while compensating the delay variation, introduces a fixed delay. This, as a minimum, should correspond to the peak PDV, thus, contributing to an increase in the overall end-to-end delay.

For fronthaul transport networks the overall delay must be limited to less than 100 microseconds over the multi-hop paths from the RRH to the BBU [7]. As a consequence, the PDV must be limited at values that can be compensated by the playout buffer within the available delay budget. If the PDV in the nodes is small, the delay added by the playout buffer may be negligible compared to the other delay components along the path. In such a case, the size of the playout buffer may be fixed within a defined maximum number of nodes in the network path. This may simplify and ease network planning and deployment.

The delay constraints need to be met both for traffic traversing a multi-hop path and for any fronthaul traffic added at any node within the transport network to a bypassing traffic stream (BP). The BP traffic, coming from the network, is a fronthaul traffic already impacted by PDV added by packet nodes (e.g. Ethernet switches) in the path. The BP traffic is aggregated with traffic generated by an

RRU (ADD) to be transferred to a BBU. Performing this aggregation with statistical multiplexing introduces a variable delay, potentially violating the required delay constraints. This requires the adoption of Ethernet switches enhanced with mechanisms for bounding both delay and PDV.

IV. BOUNDED DELAY MECHANISMS: THE TIME-WINDOW APPROACH

This section describes how one or more input "ADD" packet-streams, collected in one or more input queues of a switch, may be added with bounded delay onto an incoming bounded delay packet stream which passes through the node, referred to as a delay-bounded bypassing stream, "BP". The approach does not rely on compatible nodes (as e.g. IEEE 802.1Qbu preemption), is asynchronous and is reusing the principle of the priority mechanism from IHON, but with the goal of aggregating streams of equal priority. The target is to minimize or even avoid the PDV added to the BP stream, while minimizing and bounding the delay of the ADD stream. A time-window (TW) based mechanism is first considered, where, by relating it to the IHON TW scheme, the BP is identified as the bypassing GST class and the ADD is identified as the SM class. This mechanism sets requirements to the occurrence of sufficiently large gaps in the BP for enabling a bounded delay on the ADD. Then, a second mechanism is proposed, the Time-Window with Timeout (TWT) mechanism, which does not set requirements to traffic patterns and enables bounded delay and PDV both on the BP and the ADD.

A. ADD Insert Using TW

By using a time-window as described in [9], packets from one or more input-queues may be inserted in a gap of suitable size in a bypassing packet stream enabling a fixed delay (zero PDV) on the bypassing packet stream. When inserting ADD in vacant gaps, it can be arranged that there will be a gap of suitable size for a defined amount of data to be inserted. The data may consist of a single packet, a burst of packets or a collection of packets and gaps. The packets to be inserted, variable or fixed sizes, can be aggregated through statistical multiplexing or can be occurring from a single source (like a fronthaul RRH) or from a static multiplex of packet sources (like a fronthaul gateway aggregating traffic from multiple RRH). The principle is illustrated in Fig. 3.
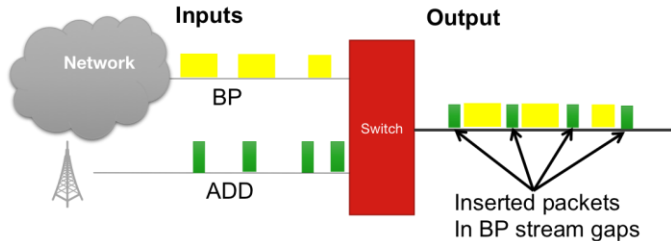

Fig. 3. TW based insertion of ADD packets in idle inter-packet gaps of a BP packet stream.

ADD packets wait for a gap of sufficient size (the time window TW) and, as a consequence, experience a variable waiting time. If gaps of sufficient size are temporarily unavailable in the BP, the traffic to be inserted will stay in the queue until gaps become available.

B. ADD Insert Using Time-Window with Timeout (TWT)

However, even if there are pre-planned gaps in a BP bypassing traffic stream, PDV imposed by the physical layer of intermediate nodes in the traversed network path may shorten these gaps, making them too small for the ADD packets to be inserted. To avoid this, the condition for inserting packets may be changed from a fixed idle time gap, equal to the ADD packet's service time, to a shorter one, or even insertion when there is no gap present. The insertion, consequently, introduces some tolerance to the PDV of the BP traffic.

An example where the gap size is smaller than the size of the data to be inserted is illustrated in Fig. 4. In case this gap is used for insertion of ADD traffic, the BP is delayed, and a certain amount of PDV is added to BP. A methodology to insert ADD traffic ensuring a balance where both the BP and ADD traffic experience a bounded delay and PDV, is to introduce a timeout mechanism so that ADD packets will always be inserted if they have not found any suitable TW within a specified maximum time-interval, the timeout T_o. Hence, if no suitable TW is found in the BP traffic within T_o, the ADD burst is scheduled for transmission in e.g. the next coming TW that is after the transmission of the current BP burst is completed. This mechanism is here referred to as Time-Window with Timeout (TWT) and an example is illustrated in Fig. 4. In practice, if an idle time-window (gap) of pre-defined size according to the ADD burst requirements (TW_min) is not occurring within a pre-defined time (T_o) the ADD burst will be given priority and inserted in the BP in the next available gap (TW_available).
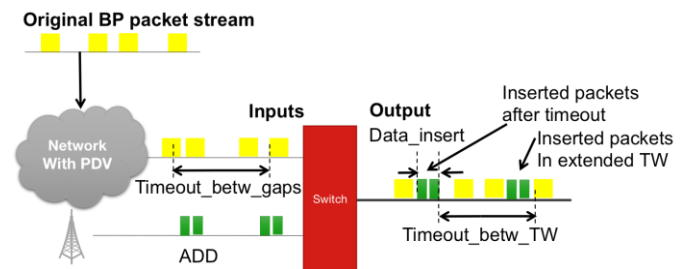

Fig. 4. Forced insertion of ADD packets into the BP packet stream by extension of the TW to the sum of the durations of the inserted packets.

However, as can be seen from Fig. 4, if the TW is smaller than the duration of packets and gaps being inserted, the result is an induced PDV on the BP. Hence, for minimizing PDV and delay on both the ADD and BP streams, a suitable T_o should be found.

The applied algorithm is described in Fig. 5. The insertion is forced when T_o expires, thus, causing insertion of the ADD immediately after the end of transmission of the current BP burst.
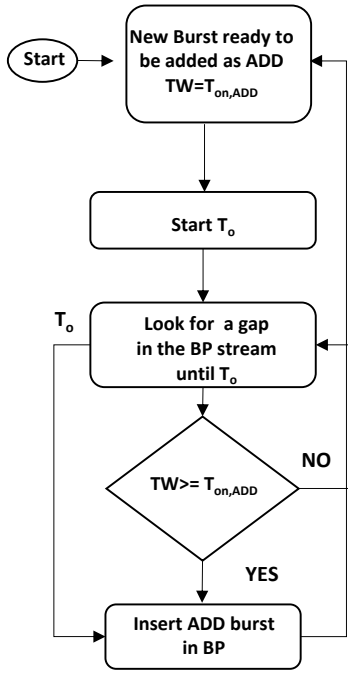
Fig. 5. Flow diagram of the TWT algorithm.

This mechanism introduces PDV on both traffic, PDV_{ADD} and PDV_{BP}, respectively. In fact, BP is affected by the forced insertion of ADD traffic when the timeout expires and ADD is influenced by the time needed to find a suitable TW, which can be variable. Hence, the introduction of the timeout allows bounding both the delay of the ADD traffic and of the BP. The trade-off that will make the reciprocal influence as much balanced as possible motivates our simulation analysis.

V. SIMULATION RESULTS

To evaluate and compare the performance of the above different mechanisms, a Bounded Delay Multiplexer (BDM) is considered, as shown in Fig. 6, representing the aggregation function implemented in an Ethernet switch. Traffic coming from N antennas on N 10 Gbit/s channels is aggregated on a 100Gbit/s channel and added to a bypassing traffic that is assumed to be generated by peripheral antennas interconnected and being shaped by the network. The objective is to evaluate the reciprocal impact of the BP and ADD traffic, and the resulting traffic's delay and PDV on the output link, induced by the proposed mechanisms.
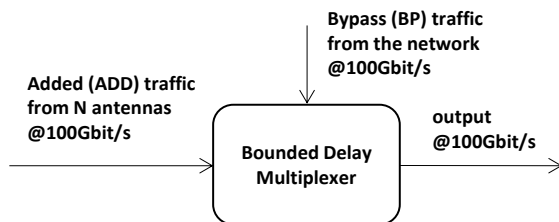

Fig. 6. Reference scheme used in simulations

The following assumptions on input traffic are made. The ADD traffic is assumed as ON/OFF with constant ON periods corresponding to one or more Maximum Transmission Unit (MTU) sized packets generated by the antennas. MTU is 1522 byte, 1518 byte with a 4 byte VLAN tag emulating eCPRI over Ethernet [3], i.e. fronthaul streams are tagged for differentiating the fronthaul traffic classes. The MTU transmission time at 100 Gbit/s is then T_{MTU}= 121.76 ns. The OFF periods are constant, and by adjusting their duration the ADD load is adjusted. The ON periods are assumed to be 3 or 6 T_{MTU}, with a corresponding ON time $T_{on,add} = 3$ T_{MTU} or 6 T_{MTU}. The ADD load is varied, with the maximum load set to $\rho_{add} = 0.5$.

To avoid random distribution of gaps between packets in the Bypass Traffic (BP), packets are added as soon as a suitable gap is found in the BP, creating bursts of packets in the BP. Hence, as a result the gaps are not randomly distributed, but occurs as bursts of gaps between bursts of packets. The bypass traffic (BP) is ON/OFF with constant ON periods emulating aggregated flows from peripheral antennas. For analyzing the impact of the added traffic, the ON period is assumed to consist of a constant number M=5 MTUs (5x1522B). This emulates a burst of added packets originating from several eCPRI sources. Therefore, packets are added as soon as possible after a BP packet. A PDV is assumed added by the network on the BP traffic, resulting in a shaping effect when traversing the network. For representing the PDV, a negative exponential distribution of the OFF periods is assumed. The physical layer induced PDV is challenging to characterize. Research either based on measurement [19], [20] or theoretical work [21], has shown that while in the core networks the traffic exhibits non-Poisson behavior, in access parts of the network, after a few levels of aggregations, the arrival process behaves as Poisson. Thus, a negative exponential distribution of the packet delay is applicable for the generating function in traffic emulators. Thus, the negative exponential distribution has been used in the simulations for emulating the delay variation induced by the physical layer. The resulting bypass traffic load is set to $\rho_{bp} = 0.5$.

As a consequence of the above assumptions, the ADD traffic is initially not affected by PDV (PDV_{ADD}=0) while the BP traffic has a PDV_{BP} represented by the exponential distribution of the OFF periods. This PDV can be thought as the result of addition of network variable delay on an original flow with the same load $\rho_{bp} = 0.5$ and constant ON and OFF periods both equal to 5 T_{MTU}. The packet delay variation of the BP traffic, indicated with $PDV_{BP,}$ is calculated for each OFF period generated by the simulator by taking the difference with respect to the original traffic with constant OFF periods (enabling gaps for ADD insertion) equal to 5 T_{MTU}.

Results are obtained by varying the value of the total offered load $\rho_{TOT} = \rho_{bp} + \rho_{add}$ by varying ρ_{add}, the number of MTUs during the ON period of the ADD traffic, and the value of the time out T_o during which the burst to be added may wait for a suitable TW. The delay D introduced by the aggregation process, the resulting PDV_{BP} and PDV_{ADD}, as well as the output load ρ_{TOT} are then evaluated by an event-driven simulator, specifically developed using the C

language. Results are taken as an average of 60 simulations for each evaluation, with a confidence interval less than 8% at 95% confidence level, for evaluations up to $\rho_{TOT} = 0.95$. The characteristics of the basic TW system without using a timeout T_o is illustrated as a reference in Fig. 7, showing the maximum and average delay introduced by the BDM on BP and ADD traffic, for $T_{ON,ADD} = 3$ T_{MTU} or 6 T_{MTU} when $T_o = \infty$. With these parameters an ADD burst waits until a sufficiently large gap is found for containing the burst (3 or 6 T_{MTU}). The BP traffic then shows a delay caused by the BDM, corresponding to a constant time equaling the duration of 3 or 6 T_{MTU}, that is 366 ns and 731 ns, respectively. The ADD traffic has a maximum delay that is above 10 microseconds even at lower load $\rho_{TOT} = 0.6$. The ADD traffic delay suddenly increases at loads ρ_{TOT} that depend on T_{ON} size, approximately at 0.75 and 0.85, for $T_{ON}=3$ and 6 T_{MTU}, respectively, as the load ρ_{add} increases with ρ_{TOT}. This means that this basic TW mechanism is effective only to limit the BP traffic delay, as expected. Instead, just waiting for a suitable time-window is not sufficient for bounding the delay of the ADD traffic. A possible solution using the TWT with timeout is therefore explored.
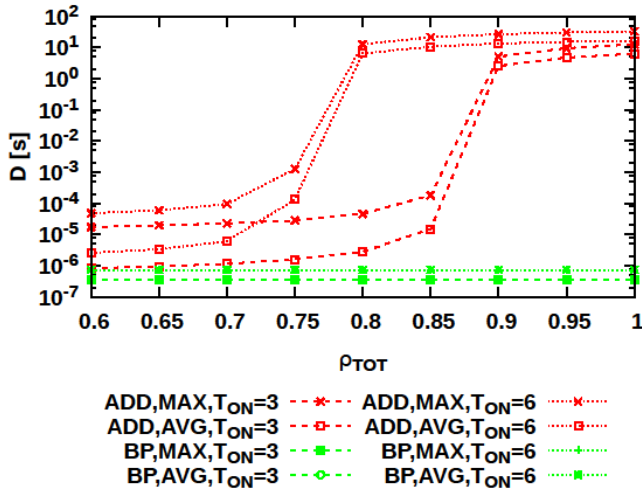

Fig. 7. Average and maximum delay of ADD and BP traffic when $T_o = \infty$, as a function of the total offered load ρ_{TOT}, varying the load of the ADD traffic ($T_{ON,ADD}$), while $\rho_{bp}=0.5$.

As a limiting case, the ADD traffic can be immediately inserted in the BP, that is a timeout $T_o=0$ is assumed, that will also be a reference for further evaluations. Figure 8 shows the delay analysis with $T_o=0$, corresponding to a strict priority of the ADD traffic, assuming $T_{ON,ADD}=6$ T_{MTU}. In this case, the maximum delay of the ADD traffic is bounded to around 608 nanoseconds, that is equal to 5 T_{MTU}. Instead BP shows a delay bound 21 µs with $\rho_{TOT}=0.95$. In addition, Fig. 9 shows the max PDV of the ADD and BP traffic at the BDM output and the contribution given by the BDM. As can be seen, the maximum PDV_{BP} of the BP traffic on output is limited to values below 12 microseconds and is even reduced at high values $\rho_{TOT}=0.95$, that is when the ADD traffic starts saturating the output channel. When the traffic on the output channel is close to the saturation, there is always ADD traffic to be inserted into the BP traffic soon

after the completion of a BP packet transmission. By forcing ADD packets even in small gaps, the BP flow becomes more regular, hence its PDV tends to drop. Going into a more detailed analysis, Fig. 10 shows the average PDV_{BP} that varies up to a maximum of 484 ns. These low average values are explained by Fig. 11 reporting the complementary cumulative distribution function of the PDV_{BP}, outlining that most values are below 2 µs, so that the resulting average values shown in Fig. 10 are below 500 ns for any value of the load. Hence, for this case ($T_o=0$), the delay and PDV of the ADD traffic are only slightly affected by the mechanism, caused by the time needed by an ADD burst to wait for a BP burst finishing its transmission on the channel.
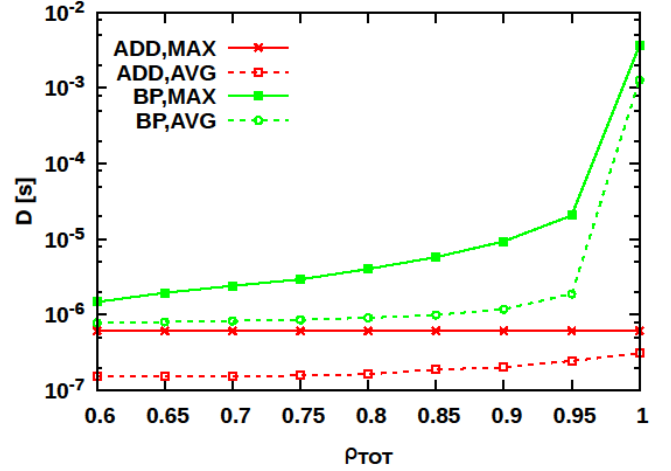

Fig. 8. Average and maximum delay of ADD and BP traffic when $T_o= 0$, as a function of the total offered load ρ_{TOT}, with $T_{ON,ADD} = 6$ T_{MTU}, $\rho_{bp}=0.5$.
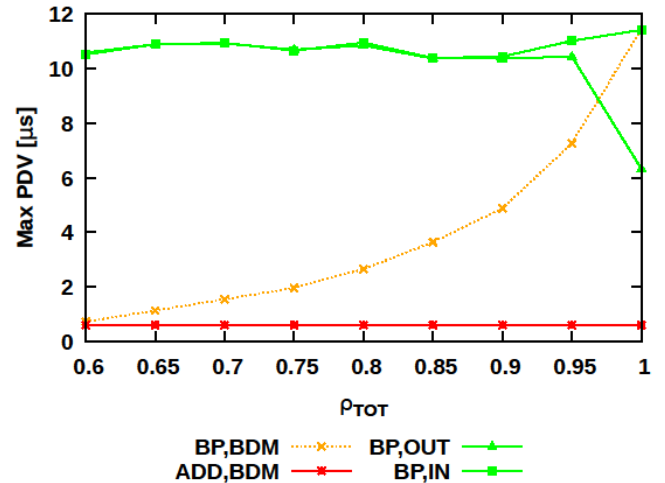

Fig. 9. Maximum PDV of ADD and BP traffic when $T_o= 0$, as a function of the total offered load ρ_{TOT}, with $T_{ON,ADD} = 6$ T_{MTU}, $\rho_{bp}=0.5$.

For mitigating the effect of the ADD traffic on the delay of the BP traffic, and potentially further compensate the network PDV on the BP traffic, different timeout values of $T_o = 3,6,9$ T_{MTU}, are considered. The objectives are to enable a slower delay increase with the load and to limit the maximum delay of the BP traffic. In Fig. 12 the maximum delay of the ADD traffic with $T_{ON,ADD} = 6$ T_{MTU} is bounded to 3 microseconds for any value of the timeout T_o. On the

contrary, the maximum delay of the BP still increases with the load, even though it is bounded to 10 µs up to the carried output load ρ_{TOT}=0.9. In this situation we cannot increase the T_o further if we want to maintain a low maximum delay of the ADD traffic. In fact, by increasing the value of T_o, the maximum delay value of the ADD traffic increases as well.
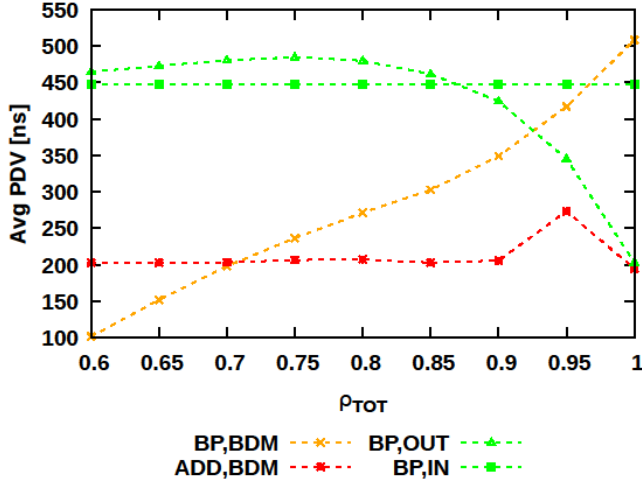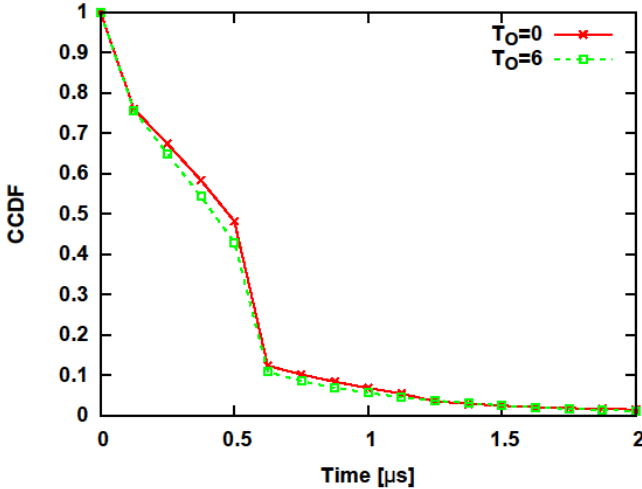

Fig. 10. Average PDV of ADD and BP traffic when T_o= 0, as a function of the total offered load ρ_{TOT}, with $T_{ON,ADD}$ = 6 T_{MTU}, ρ_{bp}=0.5.

Fig. 11. Complementary Cumulative Distribution Function (CCDF) of the PDV of BP traffic for two values of the time out, T_o= 0 and T_o= 6, with $T_{ON,ADD}$ = 6 T_{MTU}.

Fig. 13 corresponds to Fig. 12, but with $T_{ON,ADD}$= 3 T_{MTU}. Here the T_o is varied up to 5 T_{MTU}. With this value, the maximum delay of BP traffic is bounded almost to the same values as the ADD traffic for a wide range [0.6, 0.9] of the carried load ρ_{TOT}.

The maximum PDV evaluated for $T_{ON,ADD}$ = 3 T_{MTU} is plotted in Fig. 14 with the optimal value of T_o to be between T_o=3 and T_o=5, ensuring bounded delay variation for both ADD and BP traffic. The maximum PDV_{ADD} is bounded for all values of T_o to less than 2 µs. The maximum PDV_{BP} is in the range of 10 µs.
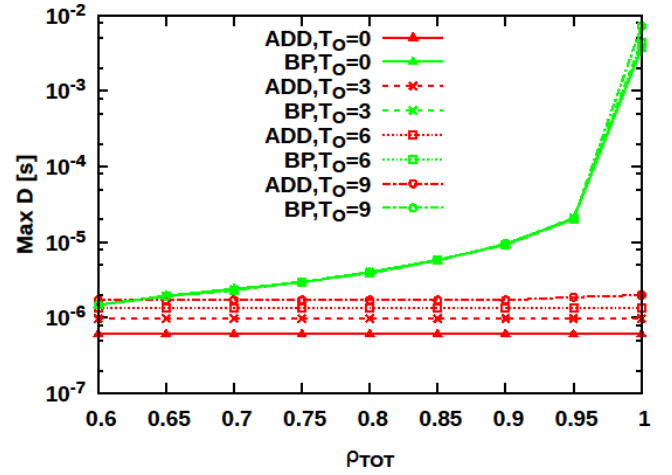

Fig. 12. Maximum delay D as a function of the total offered load ρ_{TOT}, with $T_{ON,ADD}$ = 6 T_{MTU}, varying T_o, as a parameter, ρ_{bp}=0.5.
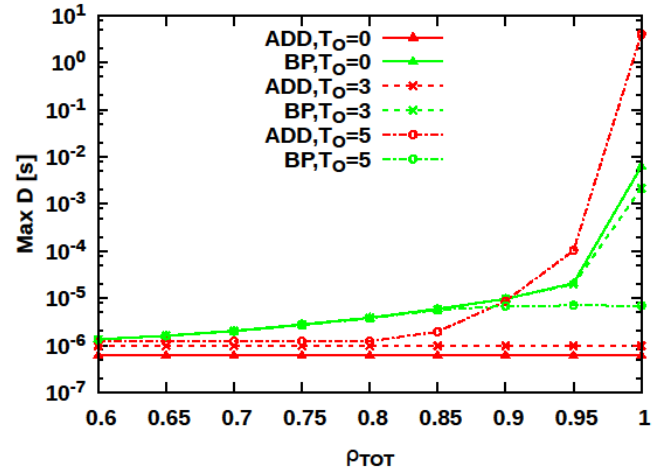

Fig. 13. Maximum delay D as a function of the total offered load ρ_{TOT}, with $T_{ON,ADD}$ = 3 T_{MTU}, varying T_o, as a parameter, ρ_{bp}=0.5.
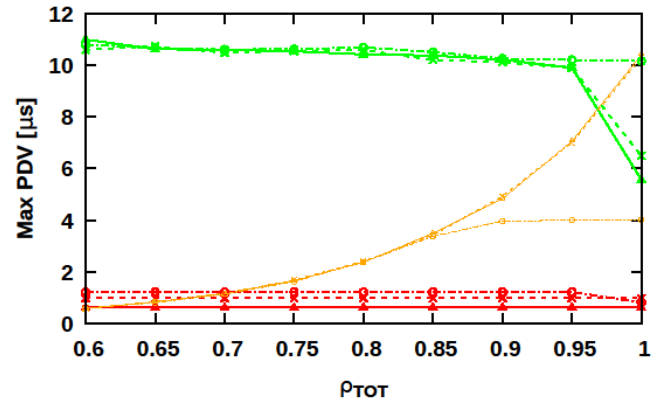

Fig. 14. Maximum PDV as a function of the total offered load ρ_{TOT}, with $T_{ON,ADD}$ = 3 T_{MTU}, varying T_o, as a parameter, ρ_{bp}=0.5.

In particular, some compensation of the PDV_{BP} may occur at values of $\rho_{TOT} > 0.95$, when the ADD traffic tends to saturate the output channel and T_o is sufficiently low for forcing insertion of ADD packets into small gaps. Similarly to figure 9, the PDV of BP traffic on the output channel is shown to be reduced with respect to the original BP flow.

In Fig. 15 the average delay is plotted in the same situation of Fig. 13 to show that for both ADD and BP traffic we have almost the same values for output traffic.
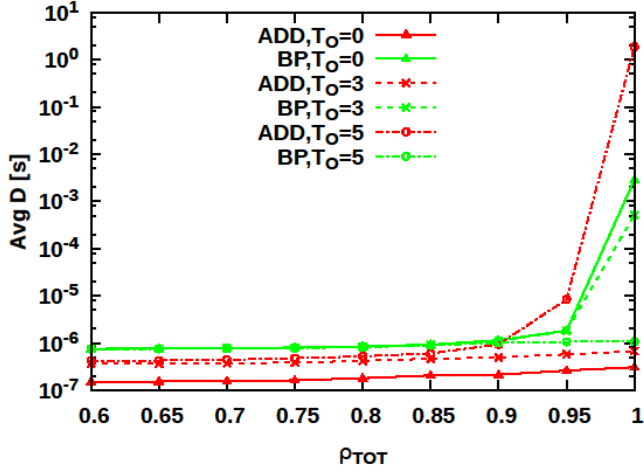

Fig. 15. Average delay D as a function of the total offered load ρ_{TOT}, with $T_{ON,ADD} = 3$ T_{MTU}, varying T_o, as a parameter, $\rho_{bp}=0.5$.

Figure 16 shows the average PDV for BP and ADD traffic as a function of the carried load on output ρ_{TOT} varying the time out T_o. We can observe that by carefully choosing the value of the timeout, at least equal in this case to 3 T_{MTU}, the average PDV of both ADD and BP traffic is approximately similar at the output of the BDM up to values of the carried load $\rho_{TOT}=0.8$, and in any case limited to 450 ns for any value of ρ_{TOT}. This interesting observation indicates that by introducing some additional bounded delay on the aggregated flows, delay and PDV characteristics may converge, enabling a balance between the two flows.
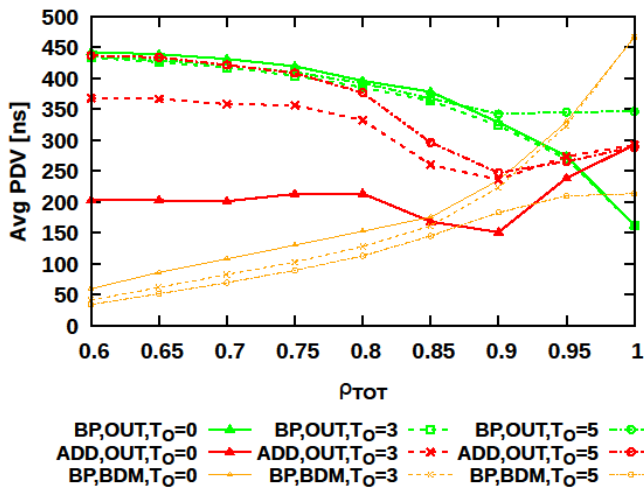

Fig. 16. Average PDV as a function of the total offered load ρ_{TOT}, with $T_{ON,ADD} = 3$ T_{MTU}, varying T_o, as a parameter, $\rho_{bp}=0.5$.

VI. EXPERIMENTAL RESULTS

An experiment was set up to demonstrate the illustrated concept of adding traffic to a bypassing traffic stream when using TW under the influence of the PDV from the physical layer. The TW without timeout mechanism, i.e. $T_o= \infty$, is implemented by TransPacket in an FPGA on Xilinx VCU110 evaluation boards, illustrated in Fig. 17(a). The implemented node has 2 x 100GE interfaces, one is used as input for the BP flow, and the other one is looped, acting as a transport path for the aggregated traffic, i.e. both ADD and BP streams are present. Furthermore, 5 x 10GE interfaces were used for adding the ADD traffic in the idle time gaps (TWs) of the BP.
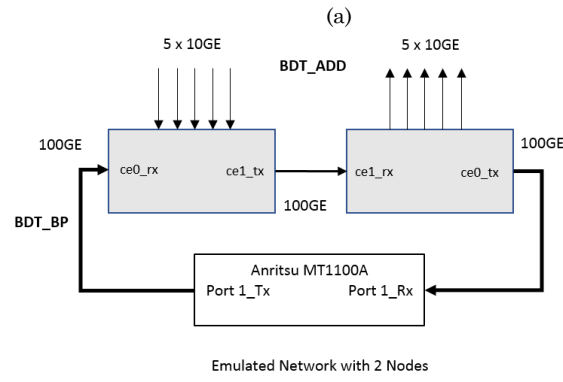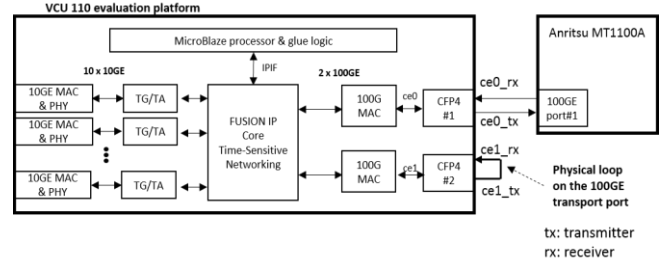

(a)

(b)

Fig. 17. (a) Experimental set up with one Xilinx VCU110 board including the FUSION IP Core from TransPacket implementing IHON principles, and Anritsu MT1100A tester. (b) Emulated network with two nodes: one 100GE BDT_BP stream and 5 x 10GE BDT_ADD inserted at the first node and dropped on the second to be measured.

An Anritsu MT11000A tester with 100ns resolution was used for generating the BP flow, while the ADD flows were generated and measured by internal traffic generators/analyzers implemented on each 10GE port in the FPGA. The transport 100GE port was physically looped so that all traffic will be looped back to be measured at the source 10GE ports for ADD, and the 100GE port for BP, respectively. Thus, the experimental setup emulates a two-node network as illustrated in Fig. 17(b).

In a first test case, a BP flow is generated with $T_{ON,BP}=5$ T_{MTU} and $\rho_{bp}=0.5$. ADD flows are generated at 10 Gbit/s at $T_{ON,ADD}=T_{MTU}$ and $\rho_{add}=0.1$, i.e. full load on the 10GE interface with an Ethernet minimum inter-frame gap (IFG) of 12 Bytes. For varying the total offered load on the transport path from 0.6 to 1, the streams are added one by one with direction towards the looped 100GE interface.
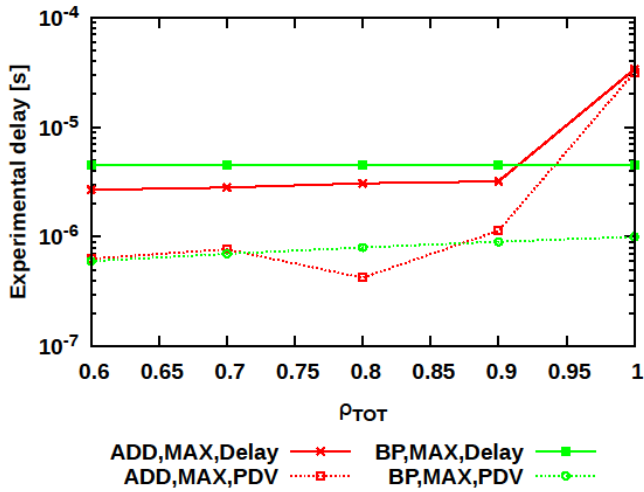
Fig. 18. Experimental Delay and PDV for the ADD and BP, as a function of the total offered load ρ_{TOT}, with $T_{ON,ADD} = 1\ T_{MTU}$, $T_o= \infty$, $T_{ON,BP} =5\ T_{MTU}$, $\rho_{bp}=0.5$.
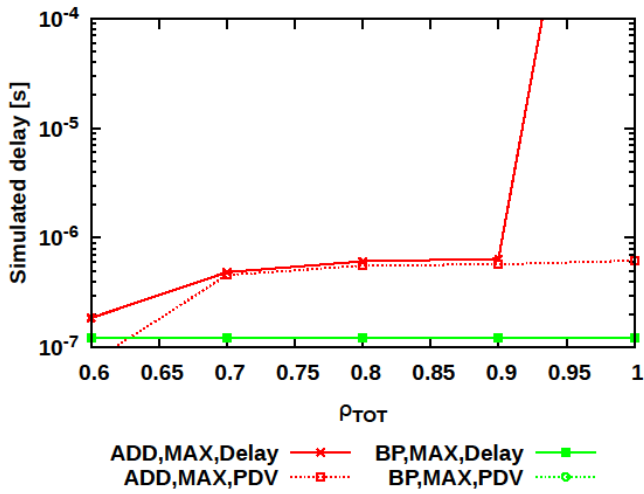

Fig. 19. Simulation results for Delay and PDV for the ADD and BP, as a function of the total offered load ρ_{TOT}, with $T_{ON,ADD} = 1\ T_{MTU}$, $T_o= \infty$, $T_{ON,BP} =5\ T_{MTU}$, $\rho_{bp}=0.5$.

Delay and PDV are evaluated for both ADD and BP and shown in Fig. 18. The end-to-end delay for BP was measured at a maximum of 4.5 µs with peak PDV 100 ns, i.e. consistently bounded through all measurements. The BP results include a two-node path, where the nodes' time-window for ADD insertion is set to a maximum packet length of 1550B that is equal to 0.19 µs, store-and-forward or serialization delay of 0.61 µs for the burst size of 5 T_{MTU}, plus the processing/pipelining delays in the IP Core set to 0.9 µs in this test case. In addition, the 100GE Xilinx board MAC/PHY reference delay and PDV measurements have shown a maximum delay of 0.45 µs per node and a PDV of up to 31 ns, while the Anritsu tester adds a maximum delay of 0.2 µs. The ADD delay is measured to 2.7 µs, i.e. it is lower than BP until the saturation point. The reason is that first the store-and-forward/serialization delay is lower as it is for one single MTU packet, second it does not undergo the fixed delay required in the BDM but only experiences a queuing delay. Furthermore, it is only a one hop

measurement, while the scenario for BP is two hops. Results demonstrate that the PDV for ADD traffic increases and is dependent on the traffic load, while BP has a fixed delay and PDV.

In Fig. 19, simulation results obtained in the same situation are also shown. The experiment confirms the effectiveness of the mechanism in bounding the delay and the PDV up to a value of the load around 0.9 corresponding to the simulated value. However, simulations do not take into account the contributions of the 100GE MAC on delay and PDV, which are present in the experiment.
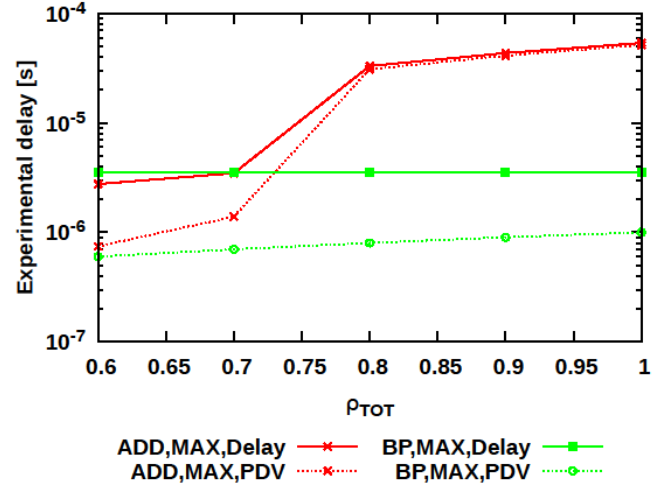

Fig. 20. Experimental Delay and PDV for the ADD and BP, as a function of the total offered load ρ_{TOT}, with $T_{ON,ADD} = 1\ T_{MTU}$, $T_o= \infty$, $T_{ON,BP} =1\ T_{MTU}$, $\rho_{bp}=0.5$

In order to evaluate/demonstrate the influence of the PDV added by the Ethernet MAC/PHYs in the gap distribution another test case was performed. The physical implementation of the Ethernet MAC influences the PDV and is important as it leads to higher delays for ADD and thus outlines the importance of the studied schemes for further bounding the delay and PDV of the ADD traffic. The second test case was evaluated through the experimental testbed where the same offered load of BP at $\rho_{bp} = 0.5$ was created by a single MTU ON and OFF periods (i.e. pre-planning the insertion of the ADD in the empty gaps of 1522B + 12B minimum IFG). Results, illustrated in Fig. 20, demonstrate that the single-packet gap distribution of the bandwidth available to the ADD traffic now becomes more limited by the PDV influence than in the first test case. The PDV at the 100GE MAC/PHY was measured at 31ns, equal to the duration of 387 Bytes. Thus, the influence on the available gap for 1 MTU of 1522Byte is higher than the traffic pattern with a gap of 5 MTU. As a result, the ADD traffic reaches the saturation point earlier, between loads of 0.7 and 0.8.

VII. CONCLUSION

Motivated by the 5G networks' need for low and deterministic delay to serve time-critical applications and fronthaul transport, we have proposed novel mechanisms for bounding network delay/PDV and explored these for an Ethernet switched network. Traffic has been added to a traffic-stream bypassing a node, using a time-window based

approach enabling a fixed delay on the bypass stream. Experiments demonstrate the practical viability of the approach of adding traffic in time windows in a bypassing traffic stream. However, PDV potentially induced on bypass traffic due to e.g. imperfect Ethernet MAC design decreases the time-window size. As a result, PDV and delay on the added traffic will be increased. Hence, as a counteraction, for balancing the PDV and delay between the two streams, a timeout mechanism ensuring traffic to be added before a defined maximum time has elapsed, has been proposed. Through simulating the performance of the mechanisms and exploring a large parameter space, main results show that a suitable choice of the timeout in relation to the traffic pattern limits and balances PDV and delay for the two aggregated traffic streams. This has been shown to be valid even for high total carried load.

Our work motivates a further study, extending the functionality of the proposed mechanisms. The goals should then be exploring how the PDV in the bypassing stream can be compensated more efficiently, extending investigations to a larger variety of traffic patterns and configurations.

REFERENCES

[1] A. Pizzinat, P. Chanclou, F. Saliou, and T. Diallo, "Things You Should Know About Fronthaul," *IEEE/OSA J. Lightwave Technol.*, vol. 33, pp. 1077-1083, 2015.

[2] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. Berger, and L. Dittmann, "Cloud RAN for Mobile Networks—A Technology Overview," *IEEE Comm. Surveys & Tutorials*, vol. 17, no. 1, pp. 405-426, 2015.

[3] Common Public Radio Interface, eCPRI Interface Specification V1.2, June 2018.

[4] IEEE Std. P1914.1, "Standard for Packet-based Fronthaul Transport Networks"

[5] IEEE Std. 1914.3-2018, "Radio Over Ethernet Encapsulations and Mappings",

[6] xRAN Fronthaul Working Group, "Control, User and Synchronization Plan Specification Version 2.0," July 2018.

[7] IEEE Std. 802.1CM, "Time Sensitive Networking for fronthaul", May 2018.

[8] D. Chitimalla, K. Kondepu, L. Valcarenghi, M. Tornatore, and B. Mukherjee, "5G Fronthaul–Latency and Jitter Studies of CPRI Over Ethernet," *J. of Opt. Comm. and Net.,* vol. 9, no. 2, pp. 172-182, Feb. 2017.

[9] R. Veisllari, S. Bjornstad, J.P. Braute, K. Bozorgebrahimi, and C. Raffaelli, "Field-trial demonstration of cost efficient sub-wavelength service through integrated packet/circuit hybrid network," *J. of Opt. Comm. and Net.*, vol.7, no.3, pp. A379-A387, 2015.

[10] R. Veisllari, S. Bjornstad, M. Fontaine, and C. Raffaelli, "High-Speed Transport and Aggregation for Ethernet Fronthaul with Low and Bounded Delay," *IEEE/OSA OFC 2019 Conf.*, Invited paper, submitted for publication.

[11] G. S. Zervas, J. Triay, N. Amaya, Y. Qin, C. Cervelló-Pastor, and D. Simeonidou, "Time Shared Optical Network (TSON): a novel metro architecture for flexible multi-granular services" *Optics Express,* vol. 19, no. 26, pp. B509-B514, 2011.

[12] A. F. Beldachi, E. Hugues-Salas, A. Tzanakaki, Y. Yan, R. Nejabati, and D. Simeonidou, "Experimental Demonstration of 5G Fronthaul and Backhaul Convergence based on FPGA-based Active Optical Transport", *Proc. IEEE ECOC 2019 Conf.*, Rome, September 2019

[13] W. Lautenschlaeger, N. Benzaoui, F. Buchali, L. Dembeck, R. Dischler, B. Franz, U. Gebhard, J. Milbrandt, Y. Pointurier, D. Roesener, L. Schmalen, and A. Leven, "Optical Ethernet - Flexible Optical Metro Networks", *IEEE/OSA J. Lightwave Technol.*, vol. 35, no. 12, pp. 2346-2357, 2017.

[14] S. Bjørnstad, D. Chen, and R. Veisllari, "Handling Delay in 5G Ethernet Mobile Fronthaul Networks," *Proc. IEEE EUCNC 2018 Conf.*, Ljubljiana, May 2018.

[15] IEEE Std. P802.1Qbu, "Standard for Local and Metropolitan Area Networks-Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks - Amendment: Frame Preemption", Jul. 2015.

[16] IEEE Std. 802.1Qbv, "Standard for Local and Metropolitan Area Networks-Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks Amendment: Enhancements for Scheduled Traffic", Mar. 2016

[17] IEEE Std. 1588-2008, "Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", Jul. 2008.

[18] IEEE Std. P802.1Qcr, "Standard for Local and Metropolitan Area Networks - Bridges and Bridged Networks Amendment: Asynchronous Traffic Shaping".

[19] A.M. Sukhov, M. Astrakhantseva, A. Pervitsky, S. Boldyrev, and A. Bukatov, "Generating a function for network delay," *J. of High Speed Net.*, vol. 22, no. 4, pp. 321-333, 2016.

[20] H. Dbira, A. Girard, and B. Sans, "Calculation of packet jitter for non-Poisson traffic," *Annals of Telecommunications*, vol. 71, no. 5–6, pp. 223–237, June 2016.

[21] G. Perez, J. Hernandez, and D. Larrabeiti, "Fronthaul Network Modeling and Dimensioning Meeting Ultra-Low Latency Requirements for 5G," *J. of Opt. Comm. and Net.*, vol. 10, no. 6, pp. 573-581, June 2018.

Steinar Bjornstad is Strategic competence and research manager at Tampnet Stavanger, Norway, adjunct associate professor IIK/NTNU Trondheim, Norway, Senior research scientist SIMULA Oslo MET, Norway. He has a master in physics from the University of Oslo (1991) and Ph.D. in telecommunication from the Norwegian University of Science and Technology (NTNU) Trondheim, Norway (2004). He was involved in multiple EU projects during his work at Telenor R&D (senior scientist, 10-year period) doing research on optical networks. He has been responsible for development of Video over IP products at Network Electronics (now Nevion) before joining Ericsson Norway, responsible for system design of Ethernet switches. From 2004 he has also been adjunct associate professor at NTNU addressing optical networking topics. He is former CEO and CTO in TransPacket which he founded in 2009. Since 2015 he has been involved in IEEE 802.1 Ethernet and 1914.1/1914.3 next generation fronthaul interface/radio over Ethernet standardization. He is author/co-author of more than 60 scientific papers and has several international patent-families.

Raimena Veisllari is system architect and product manager at TransPacket, focused in research and design of time-sensitive Ethernet mechanisms with special emphasis on new application areas related to 5G mobile transport and optical packet/circuit networks. She holds a joint master's degree in security and mobile computing from KTH, Sweden and NTNU, Norway (2010), and a PhD in Telematics from NTNU (2016). She has more than 30 papers in scientific conferences and journals, is an invited guest lecturer at the University of Bologna and invited speaker at OFC 2019. Currently she is involved in the research project 5G-PICTURE supported by the European Horizon 2020 initiative.

David Chen Dr. David Chen has more than twenty years of industrial experiences in advanced technologies and the wireless industry. He is currently the Senior Architect at Nokia with a focus on 5G & Cloud RAN transport architecture and X-haul Proof of Concept product design. Dr. Chen has been actively participating and contributing in standardization bodies, such as IEEE802.16,

802.1/.3, 1588, 1904/1914, and CPRI. He has also been the invited speaker in various industrial forums such as 4G World, IWPC, ITSF, WSTS, ICC and many more. He was the merger and acquisition technical advisor and worked closely with the venture capital branch to recommend and evaluate multiple innovative wireless backhaul startup companies, many of which led to successful acquisitions. He has numerous patents and papers in various journals, conferences, symposiums, and standard bodies. He received his bachelor's degree in Electrical Engineering from National Taiwan University and both his Master degree and PhD in Electrical Engineering from Northwestern University (USA).

Federico Tonini (M'2017) Federico Tonini received his B.S. degree in Electronics and Telecommunications Engineering and the M.S. degree in Telecommunication Engineering from University of Bologna in 2012 and 2015, respectively. He is currently pursuing his Ph.D. in Electronics Engineering, Telecommunications, and Information Technology at DEI, University of Bologna. His research interests include 5G communication system, network design and optimization techniques.

Carla Raffaelli (SM'2016) is Associate Professor at the University of Bologna. She received her M.Sc. and Ph.D degrees in electronic and computer engineering (University of Bologna, Italy), in 1985 and 1990, respectively. Her research interests include performance analysis of telecommunication networks, switch architectures, optical networks and 5G networks. She actively participated in many National and International research projects, such as the EU funded ACTS-KEOPS, the IST-DAVID, the e-photon/One and BONE networks of excellence. She is author or co-author of more than 150 conference and journal papers mainly in the field of optical networking and network performance evaluation. She regularly acts as a reviewer for top international conferences and journals. She served as Technical Program Committee member in several Top International Conferences, like ICC and ONDM, and served as Technical Program Committee co-Chair in ONDM 2011. Since October 2013 she is a member of the editorial board of the journal Photonic Network Communications by Springer. She is IEEE Senior Member since 2016. She is the Director of the International Telecommunications Engineering Master's Degree at the University of Bologna, Italy, since 2015.