Scheduling in an Ethernet Fronthaul Network

M. K. Al-Hares[1], P. Assimakopoulos[1], *Member*, *IEEE*, D. Muench[2] and N. J. Gomes[1], *Senior Member*, *IEEE*
[1]Communications Research Group, University Of Kent, Canterbury, United Kingdom
[2]ADVA Optical Networking SE, Munich, Germany
(ma658@kent.ac.uk)

***Abstract*—This paper investigates and compares the performance of different scheduling techniques in an Ethernet fronthaul network in the presence of both time-sensitive/high priority and background traffic streams. A switched Ethernet architecture is used as the fronthaul section of a cloud radio access network (C-RAN) and a comparison of two scheduling schemes, strict priority scheduling and time-aware shaping, is carried out. The different streams are logically separated using virtual local area network identifiers and contend for the use of trunk links formed between aggregator/switch nodes. The scheduling schemes are applied in the access and trunk ports in the fronthaul, and need to handle the queue management and prioritization of the different streams. In such cases, contention-induced latency variation has to be characterized, especially when the fronthaul transports precision time protocol traffic, as it directly leads to errors in timestamping. OPNET models for strict priority and time-aware schedulers have been built and employed, and simulation results are used to compare the performance of the two scheduling schemes.**

***Keywords*—PTP; Synchronization; C-RAN; fronthaul; IEEE802.1Qbv; Ethernet; priority**

I. INTRODUCTION

Due to its ubiquitous and potentially low cost nature, Ethernet is being considered as the transport technology of the future centralized (or cloud) radio access network (C-RAN) [1-5]. Ethernet offers fully standardized carrier-class operation administration and management (OAM) and can lead to structural and operational convergence within an operator's network. It is agnostic to the traffic that is being encapsulated – a clear advantage – but this last characteristic, also results in the main drawback of Ethernet: Any implicit or explicit synchronization within the traffic stream that is being encapsulated is lost since Ethernet, in its native form, lacks any form of synchronization. One example is the encapsulation of the common public radio interface (CPRI) by Ethernet (*see* the radio-over-Ethernet (RoE) IEEE1914.3 standard [4]).

Current C-RAN implementations transport In-phase and Quadrature (IQ) radio samples from base station baseband units (BBU) to remote radio heads (RRH) using CPRI [6]. This centralized approach is limited in its ability to scale to bandwidth requirements that will be imposed on the transport infrastructure by long-term evolution-advanced (LTE-A) (as a result of carrier aggregation) and, even more so, by envisaged 5^{th} generation (5G) signals and massive multiple-input and multiple-output (mMIMO) applications. As a means of reducing the data rate requirements of the future fronthaul, the implementation of different physical layer functional subdivisions (or "splits") has been proposed [1-3, 7] and is also the subject of current standardization efforts [8]. Furthermore, the combination of functional splitting with the use of Ethernet technology in the fronthaul can lead to cost reductions (by leveraging the ubiquity of Ethernet equipment) and performance enhancements (mainly from the ability to obtain statistical multiplexing gains), in addition to a reduction in data rate. A further advantage that Ethernet technology can bring is its direct integration with cloudification and virtualization techniques, through the use of Ethernet switches/aggregators which will also form the new aggregation points where the statistical multiplexing gains will be obtained.

Such an implementation will need to fulfil the time and frequency synchronization (syntonization more accurately for the latter) requirements, by employing, through some modifications, existing protocols such as synchronous Ethernet (SyncE) (*see* [9] and associated standards) and the precision time protocol (PTP) [10]. However, a time-synchronized network does not solve the problem of contention in aggregation points. While contention in general can lead to a violation of key performance indicator (KPI) specifications for a given system (namely latency and latency variation), it can also detrimentally affect any "in-line" timing protocol, such as PTP.

A number of scheduling regimes are being proposed including (among others) the IEEE 802.1CM standard's (Time-Sensitive Networking for Fronthaul) [5] profile A which is based on strict priority (SP) scheduling and the IEEE 802.1Qbv (Enhancements for Scheduled Traffic) amendment [11]. The latter specifies a procedure to enable an aggregation point (a bridge in IEEE 802.1Q nomenclature) or end-point to transmit packets from a traffic class queue with reference to a known timeframe (and therefore assumes a time-synchronized network). It does this by employing a transmission gate per traffic class queue to allow only one traffic class queue to access the network at a specific time.

It is, then, important to investigate the performance of these scheduling techniques and measure the effect of each technique on the different traffic classes that may exist over the fronthaul network. A number of papers have analysed the effects of contention, with [12-14] and without [15] scheduling, in an Ethernet fronthaul and its effects on KPIs. In this work, time-aware shaper (TAS) and SP models are built and configured, respectively, in OPNET.

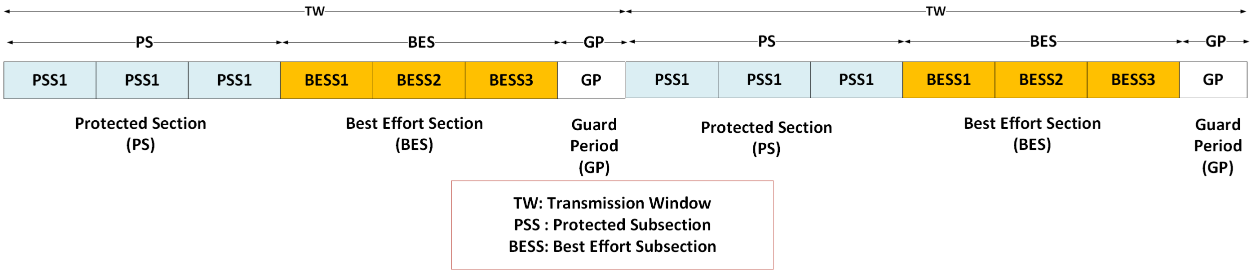

Fig. 1. Generic time window, window section and subsection plan based on IEEE 802.1Qbv.

Simulations are then carried out and are used to quantify the resulting frame-delay variation (FDV) of a PTP traffic stream when this contends with background traffic for a number of different background traffic regimes. The latter are used to represent variable-frame size traffic (e.g. traffic that would be generated by different functional splits) or constant-frame size traffic (such as control & management (C&M) or CPRI). Section II will introduce the operational principles of the TAS and SP schedulers while Section III will describe the OPNET implementation. Simulation results are then presented in Section IV and the conclusions in Section V.

II. Scheduling Operation Principle

A. *Time-aware shaper (TAS)*

The IEEE 802.1Qbv standard is introduced as a solution to contention-induced FDV in network bridges [11]. The main concept of the standard is to allow time-referenced transmissions from port queues in network bridges and traffic sources based on explicit time scheduling. Thus, a prerequisite to implementing this time shaping mechanism is to have an overlaid time synchronization network. The TAS will assign specific window sections to each traffic and only allow the traffic to pass through a bridge during this time window section. The division of the total transmission time (i.e. encompassing all traffic sources), termed the time window (TW), into the different sections is shown in Fig. 1. High priority (HP) traffic is assigned to one TW section while the best effort traffic is assigned to a different TW section (within the total transmission time). Each window section may be formed by a number of sub sections and each section or sub-section can be allocated to one or more traffic streams. The section assigned to the HP traffic is termed the protected section (PS) while a best effort section (BES) is allocated to the lower priority traffic. To prevent the best effort traffic from overrunning into the PS, part of the transmission time after the BES will be idle and not assigned to any traffic, in effect forming a guard period (GP).

The size of the sections and the sub-sections is associated with the number of prioritized or time-sensitive traffic streams within the network. An example of high-priority traffic in the network is PTP traffic while C&M traffic will usually be treated at a lower priority setting. It is possible within the BES to assign priority levels to the different lower priority streams and employ an “intra-section” scheduler such as SP, weighted-round-robin (WRR) or weighted fair queuing (WFQ).

B. *Strict Priority (SP)*

With SP, the queue with the highest priority transmits first and then the other lower priority queues will transmit, one by one, in the order of their priority setting. Thus, the lowest priority queue has to wait for all higher priority queues to finish transmission in order to have an opportunity to start its own transmission. On the other hand, if a HP frame arrives while a lower priority queue is being served, the HP frame will have to wait for the current transmission to finish. SP may be used only with extremely delay/jitter sensitive traffic such as PTP traffic, as it can lead to starvation of the lower priority queues.

III. Simulation Platform Implementation

A. *Time Aware Shaper Implementation*

The implementation of the TAS algorithm in OPNET is shown in Fig. 2. It begins by initializing, in the first instance, the TW section boundaries, for the duration of the first TW, in the Ethernet switch. The switch’s medium access control (MAC) layer receives an Ethernet frame from the physical input port, inserts it into a buffer in the input port and checks whether it is being received within the allocated section. If it is, the switch will allow the traffic to pass through, by sending it to the output port queue. If the next frame to follow is also within the allocated section, it is also allowed to pass through. If a frame is received outside its allocated section, the frame is dropped. The last check in the algorithm is whether a TW section (PS or BES) has expired. If it has, the section boundaries are updated to correspond to the next allocation (i.e. in the TW to follow). While the PS and BES sizes are specified based on the amount of transmitted traffic from each traffic type, the GP can be designed based on the serialization delay of a best effort traffic frame.

B. *Strict Priority*

Strict priority can be readily configured for use in OPNET as models already exist for it. The PTP traffic is assigned to a higher priority while the background traffic is assigned to the lower priority setting.

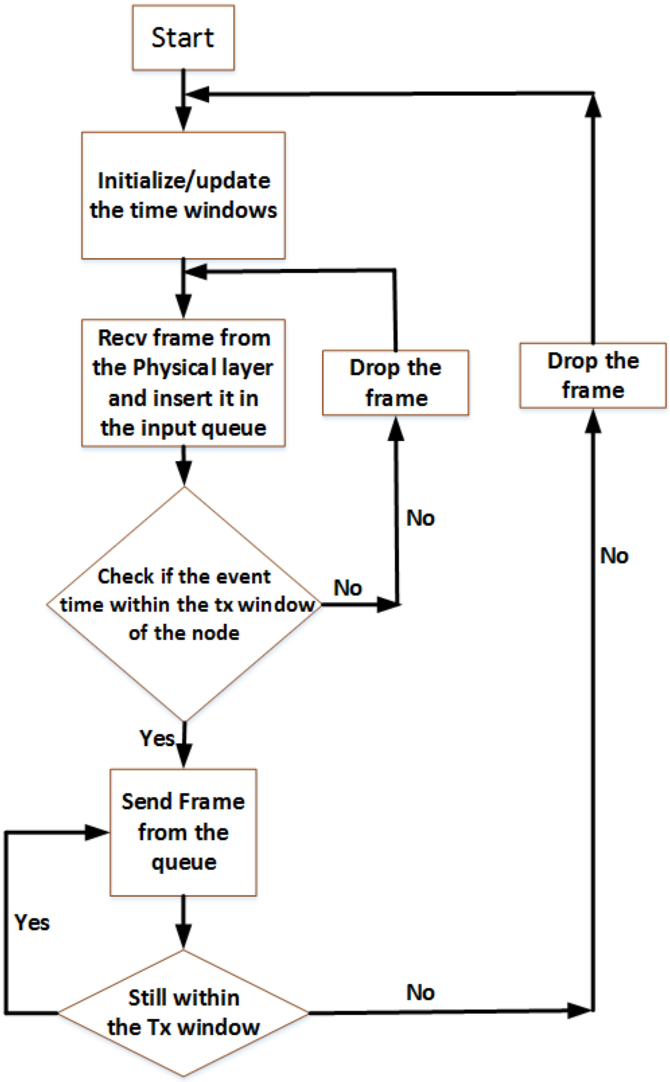

Fig. 2. Implementation Algorithm in Opnet.

C. *Network Description*

The network implementation in OPNET is shown in Fig. 3. It consists of two traffic generators (TGs); one of them represents the PTP grandmaster (TG1) while the other (TG2) is the best effort traffic generator. TG1 sends data over VLAN ID 10 while TG2 sends data over VLAN ID 20 in a port-based configuration (i.e. the end stations do not tag the frames). The distance from TG1 to the first bridge is set to 200 m while the distance from TG2 to the first bridge is set to 1700 m. The distances between the traffic generators and the first bridge are set simply as an example.

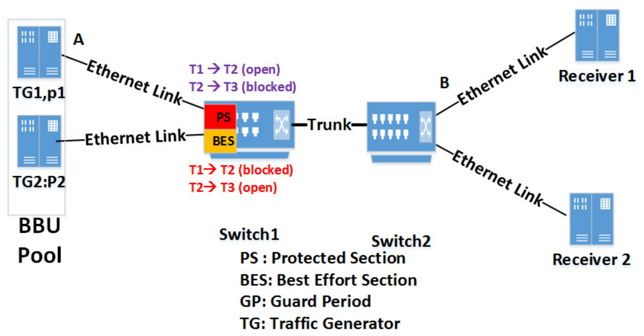

Fig. 3. Network Scenario implemented in Opnet. All network interfaces are 1 GbE.

IV. Simulation Results

To compare the queuing regimes, two different scenarios are implemented. In both scenarios, the PTP traffic (the HP traffic) and background traffic (the BE traffic) are transmitted over the same network segment (a trunk link), with the PTPv2-emulating traffic stream assigned to a higher priority setting. TG1 generates 32 timing messages per second per PTP slave. The number of slave stations is 50 (note that these are modelled through the amount of traffic generated and the corresponding utilization in the trunk and not as separate receivers) and each sync message is formed as a 68 octet frame. Note that the amount of background traffic that shares the trunk link with the PTP traffic may not correspond to the same number of receiving stations. The simulation results in this work focus on the high priority traffic (PTP traffic) toward the network edge.

A. *Constant Background Frame Size Scenario*

In this scenario, background traffic is generated as a burst of fifty frames, with an inter-frame gap of 20 µs and a frame size of 1000 octets. This traffic source may represent either CPRI-type traffic or C&M traffic. The PS duration is set to 50 µs. The GP is allowed to vary from zero to the value of the serialization delay of a 1000-octet frame. The TW size is 625 µs

Fig. 4 shows the peak and average FDV results for SP and for TAS with different GPs. The results show that the SP performance is equivalent to a TAS implementation with zero GP. Specifically, the peak and average FDV with TAS are upper-bounded to the peak and average FDV with SP. This makes sense as both schedulers cannot resolve cases where a background traffic transmission is ongoing, i.e. there is no pre-emption being employed in the network (*see* for example the SP results in [13]).

As the GP is increased, both the average and max FDV with TAS reduce steadily until they reach zero at a GP of 6 µs. This value corresponds to a serialization of a large part (75%) of a background traffic frame.

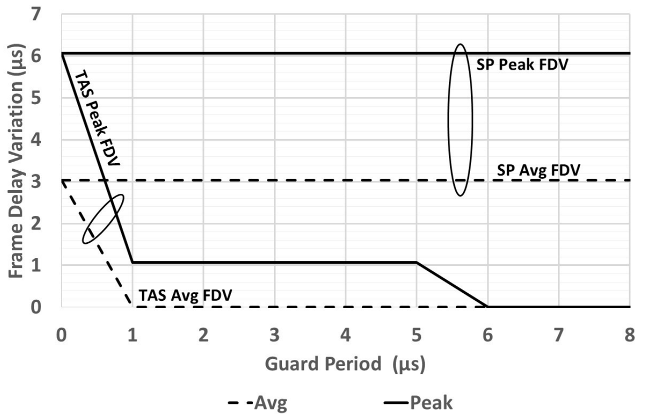

Fig. 4. Average and peak FDV for the PTP traffic with SP and TAS with different GPs. The background traffic source is constant frame-rate and constant frame size.

Note that the worst case would be a full serialization (i.e. 8 µs) but as the background source is constant frame-rate and frame size this worst case is not observed in these results due to the relative timings of the background and PTP traffic generators in the simulation. Note also that the peak-to-average ratio of FDV can be very large and obviously, this ratio will depend on the transmission pattern of the sources.

B. Variable Frame Size Scenario

This scenario is similar to the first scenario but with a varying frame size for the background traffic. The traffic source is meant to represent functional split traffic, for e.g. fifty user allocations per LTE subframe (i.e. 50 frames every 1 ms), in a MAC/PHY split (3GPP option 6 [16]). Note also that a constant (or close to constant) number of allocations could arise as a result of employing statistical multiplexing gains over a trunk link. Two different settings are used: The first follows a normal distribution with a mean value of 1000 octets and variance of 200 octets (Fig. 5). The second is similar albeit with an increased variance of 500 octets (Fig. 6).

The results show that the peak and average FDV is increased (compared to the first scenario) for both SP and TAS with zero GP and approaches the serialization delay of a full background traffic frame. Furthermore, the peak FDV for the results of Fig. 6 reaches zero at a GP that is equivalent to the serialization delay of a full background traffic frame. This is indicative of the dependence of the scheduler performance, with regards to FDV, on the transmission pattern characteristics of the traffic sources. The larger variance in the traffic pattern in this case results in the occurrence of the worst-case scenario of a background traffic frame serialized right at the end of its BES allocation.

Fig. 7 is a zoom-in of Fig. 6 in the x-axis range from 0 to 1 µs. The large peak-to-average ratio of FDV is clear in these results.

The small inset shows the resulting worst-case timestamping error with PTP for the peak FDV values shown in Fig. 7. The worst-case assumption is that this peak FDV is encountered in one direction of traffic (either downlink or uplink) while there is zero FDV in the opposite direction.

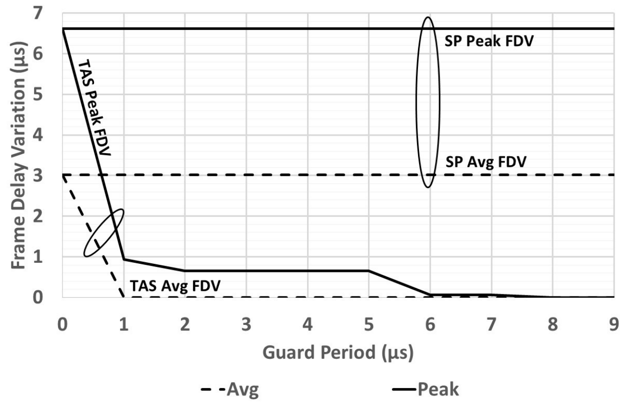

Fig. 5. Average and peak FDV for the PTP traffic with SP and TAS with different GPs. The background traffic source is constant frame-rate with a varying frame size following a normal distribution with mean of 1000 octets and variance of 200 octets.

This result shown the main limitation of SP which although can reduce significantly the average FDV, the peak FDV remains constant and can potentially result in large PTP timestamping errors (depending on the size of the background traffic frame). TAS on the other hand looks promising in its ability to reduce FDV (and thus timestamping errors) as the GP is increased, or eliminate FDV entirely when the GP is sufficient to eliminate contention. The drawback in this case is the increased end-to-end latency especially if the number of aggregation nodes becomes large.

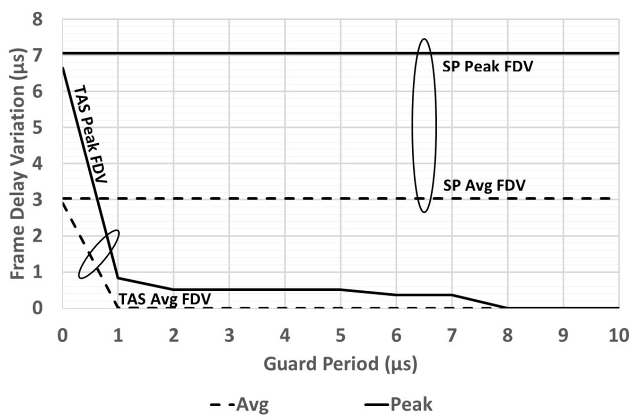

Fig. 6. Average and peak FDV for the PTP traffic with SP and TAS with different GPs. The background traffic source is constant frame-rate with a varying frame size following a normal distribution with mean of 1000 octets and variance of 500 octets.

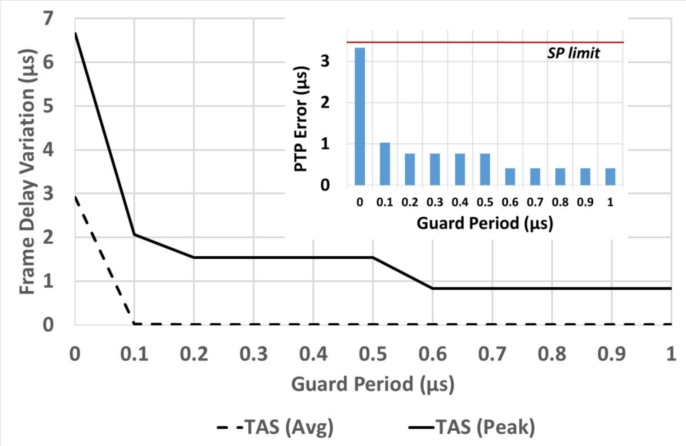

Fig. 7. Zoom-in in the region of GPs from 0 to 1 µs for the results of Fig. 6. The inset shows the worst-case PTP timestamping error that would result from the peak FDV values.

V. Conclusion

A comparison between two scheduling schemes, strict priority (SP) and time-aware shaping (TAS) has been carried out using a simulation model designed and built in OPNET. The comparison focuses on the performance of the precision time protocol (PTP), in terms of frame-delay variation (FDV), when contention with background traffic takes place in an Ethernet fronthaul. The results show that the average and peak FDV of TAS are upper-bounded to the average and peak FDV of SP. This worst-case occurs when the GP in TAS is set to zero, but as the GP is increased both average and peak FDV reduce steadily, until FDV is completely eliminated. The GP that is required to eliminate FDV has a strong dependence on the statistical variations of the traffic sources. The stronger the variation, the closer the required GP needs to be to a full background traffic frame serialization delay. The obtained peak FDV results are extrapolated to worst-case PTP time stamping errors and it is shown how these errors reduce as the GP in TAS is increased. Proper operation of PTP and the required scheduling to guarantee it will be of fundamental importance in future C-RAN fronthauls that employ Ethernet transport.

Acknowledgment

This work was carried out within the framework of the European Union's Horizon 2020 research and innovation programme under grant agreement No 644526 (iCIRRUS project) and EPSRCs "Towards an Intelligent Information Infrastructure (TI3)" programme (NIRVANA project). Mohamad Kenan Al-Hares acknowledges the funding through an EPSRC Doctoral Training Partnership (DTP) and Philippos Assimakopoulos by the NIRVANA project. The authors would like to thank Jörg-Peter Elbers for interesting discussions. Data used in this work is stored in the Kent Academic Repository (https://kar.kent.ac.uk/).

References

[1] iCIRRUS (July 2015), "D2.1: iCIRRUS intelligent C-RAN architecture," [Online]. Available: http://www.icirrus-5gnet.eu/category/deliverables/

[2] iCIRRUS (Jan. 2016), "D3.1: Verification of Ethernet as transport protocol for fronthaul / midhaul," [Online]. Available: http://www.icirrus-5gnet.eu/category/deliverables/

[3] N.J. Gomes, V. Jungnickel, P. Chanclou, J.-P. Elbers, and P. Turnbull, "A flexible, Ethernet fronthaul for 5th generation mobile and beyond (Invited),", in *Optical fiber Commun. Conf. (OFC)*, 2016, Anaheim, CA, 2016, paper W3C.1

[4] "Standard for Radio Over Ethernet Encapsulations and Mappings," IEEE Standard P1914.3 [Online]. Available: http://sites.ieee.org/sagroups□1914/p1914□3/

[5] "Time-Sensitive Networking for Fronthaul," IEEE Standard P802.1CM [Online]. Available: http://www.ieee802.org/1/pages/802.1cm.html

[6] CPRI (Oct. 2015), "CPRI Specification V7.0, Interface Specification," [Online]. Available: http://www.cpri.info/spec.html.

[7] U. Dötsch, et al., "Quantitative analysis of split base station processing and determination of advantageous features for LTE," in *Bell Labs Tech. J.*, vol. 18, no. 1, pp. 105-128, Jan. 2013

[8] "Next Generation Fronthaul Interface" IEEE 1914 Working Group [Online]. Available: http://sites.ieee.org/sagroups□1914.

[9] "Distribution of timing information through packet networks," ITU-T Recommendation G.8264,May 2014 [Online]. Available: http://www.itu.int/rec/T□REC□G.8264□201405□I.

[10] "Precision time protocol telecom profile for phase/time synchronization with full timing support from the network," ITU-T Recommendation G.8275.1, July 2014 [Online]. Available: https://www.itu.int/rec/T□REC□G.8275.1/en.

[11] "Enhancements for Scheduled Traffic," IEEE standard 802.1Qbv [Online]. Available: http://www.ieee802.org/1/pages/802.1bv.html

[12] P. Assimakopoulos, M. K. Al-Hares, S. Hill, A. Abu-Amara, and N. J. Gomes, "Statistical distribution of packet inter-arrival rates in an Ethernet fronthaul," in *IEEE Int. Conf. on Commun. Workshops (ICC)*, Kuala Lumpur, Malaysia, 2016, pp. 140-144

[13] M. K. Al-Hares, P. Assimakopoulos, S. Hill, and N. J. Gomes "The Effect of Different Queuing Regime On a Switched Ethernet Fronthaul," in *IEEE Int. Conf. on Transp .Optic.Net*. (ICTON), Trento, Italy, 2016, pp. 1-4

[14] T. Wan and P. Ashwood-Smith, "A Performance Study of CPRI over Ethernet with IEEE 802.1Qbu and 802.1Qbv Enhancements," in *Global Commun. Conf. (GLOBECOM)*, San Diego, CA, 2015, pp. 1-6

[15] P. Assimakopoulos, M.K. Al-Hares, and N.J. Gomes, "Switched Ethernet fronthaul architecture for cloud-radio access network", in *OSA/IEEE J. Optical Commun. and Netw*., vol: 8, no. 12, pp. B135-B146, Dec. 2016

[16] 3GPP, "Study on New Radio Access Technology; Radio Access Architecture and Interfaces (Release 14)," 3GPP TR 38.801 V0.4.0, Aug. 2016 [Online]. Available: http://www.3gpp.org/DynaReport/38-series.htm