Dataplane Measurements on a Fronthaul and Backhaul integrated network

Thomas Deiß[1*], Jorge Baranda[2], Luca Cominardi[3], Luis Miguel Contreras Murillo[4], Jessé Gomes[5], Sergio González[6], Paola Iovanna[7], Josep Mangues-Bafalluy[2], Nuria Molner[6,8], José Núñez-Martínez[2], Antonio de Oliva[6], Stefano Stracca[7]

[1]Nokia (Germany), [2]CTTC (Spain), [3]Interdigital (UK), [4]Telefonica (Spain), [5]HHI (Germany), [6]UC3M (Spain), [7]Ericsson (Italy), [8]IMDEA Networks Institute (Spain)

***Abstract*—Future transport networks serving next generation accesses are expected to carry both fronthaul (FH) and backhaul (BH) traffic simultaneously. This new concept of network which integrates the FH and BH traffic over the same transport substrate is called Crosshaul. A Crosshaul network will use heterogeneous technologies, such as fiber, mmWave, or microwave, and selects the most appropriates ones depending on the use case. Moreover, the softwarization/virtualization trend on the networking industry indicates that Virtual Network Functions (VNFs) will process and exchange both BH and FH data plane traffic. This paper presents performance measurements on promising technologies for the implementation of a Crosshaul network. We investigate to which extent the requirements to carry FH traffic are satisfied by mmWave links, software and multi-layer hardware switches.**

Keywords—Crosshaul, softswitch, OpenFlow, mmWave, throughput, latency, jitter

I. INTRODUCTION

A couple of networking trends are influencing the deployment of transport networks for 5G radio access networks (RAN): softwarization and network function virtualization (NFV). To implement the Cloud-Radio Access Network (C-RAN), baseband units (BBU) can be deployed as VNFs in a centralized location, requiring the transport of fronthaul (FH) traffic to sites deploying remote radio heads (RRHs). Furthermore, denser RAN deployments are expected, with smaller cells connected through different backhaul (BH) transmission technologies depending on the use case and deployment constraints. Combining these trends, transport networks for 5G will use heterogeneous transmission technologies to carry both FH and BH traffic between virtualized network functions and forwarding elements.

Softwarization of network control or software defined networks (SDN) is important for three aspects: 1) different services such as high capacity or low latency communication will be deployed, requiring a fine grain control, 2) decrease service deployment to hours instead of months as with manual approaches, 3) applications on top of the control plane to optimize the network performance, e.g. partially shutting down network elements to save energy based on traffic patterns.

FH traffic poses stringent requirements on latency, jitter, throughput, and packet loss. For instance, a 20MHz 2x2 MIMO LTE cell requires about 2.5Gbps bandwidth per direction for Common Public Radio Interface (CPRI) [1] [2] uncompressed traffic. A typical one-way latency requirement is 100µs, which corresponds to 20km over fiber, while jitter should be as small as 65ns. When carrying CPRI over a packet network, playout buffers at the end of a path can compensate jitter in the network. The use of wide carriers and massive MIMO in 5G would increase the bandwidth requirements for a functional split comparable to CPRI to a prohibitive level. Thus, new functional splits are discussed for the 5G radio stack [3]. Functional splits high in the protocol stack are expected to have characteristics like backhaul traffic; functional splits low in the protocol stack are expected to have much relaxed bandwidth, latency, and jitter requirements compared to CPRI. The traffic of such low functional splits is still considered as FH traffic. 5G use cases with low latency requirements on application level in the order of 1ms round-trip time imply as well stringent latency requirements on the transport network, independent of the RAN functional split.

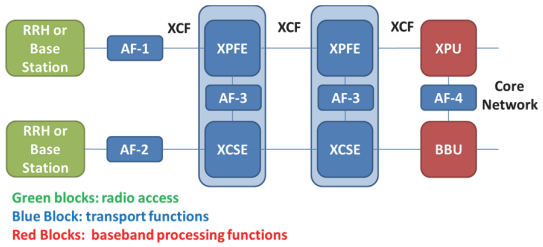

Fig. 1: 5G-Crosshaul data plane architecture [5]

A Crosshaul network is a common infrastructure for FH and BH traffic able to support different services in terms of latency and bandwidth. The reference 5G-Crosshaul architecture [4], [5] is based on three-layer nodes; packet (5G-Crosshaul Packet Forwarding Element, XPFE), circuit and optical (5G-Crosshaul Circuit Switching Element, XCSE), see Fig. 1. In the 5G-Crosshaul architecture adaptation functions (AF) are required to perform several operations. They provide en- and de-capsulation to separate traffic of different types or tenants. AFs are also used to map the different transmission technologies frames to a common framing used within the 5G-Crosshaul network. The frame is called 5G-Crosshaul Common Frame (XCF) in Fig. 1 and is implemented as a standard IEEE 802.1ah Provider Backbone Bridge (PBB) framing. Finally, AFs are also used to adapt packet flows to the circuit or optical layers. The circuit and

The authors of this paper have been sponsored in part by the project H2020-ICT-2014-2 "5G-Crosshaul": The 5G Integrated fronthaul/backhaul" (671598)
* Corresponding author email: thomas.deiss@nokia.com

optical layers allow off-loading traffic from the packet layer to provide a very low latency path.

In this paper, we present throughput, latency, jitter, and loss rate measurements on various data plane elements of the path among RRHs and BBUs. We characterize the performance of data plane elements contributing to the end-to-end performance of the 5G-Crosshaul network by analyzing i) mmWave links, ii) the adaptation among packet and circuit layers, and iii) software switches. We analyze the performance of both indoor and outdoor mmWave links and investigate the impact of indoor beamforming training procedures. The adaptation among the packet and circuit layers aims at providing constant latency, i.e. very small jitter. We present measurements of an exemplary software switch – Lagopus [6], selected for its good performance due to the use of Intel DPDK libraries. The measurements show the impact of configuring the underlying packet processing as well as the impact of the complexity of flow tables as defined in the 5G-Crosshaul OpenFlow pipeline.

The rest of this paper is structured as follows: section II describes the different test setups. Section III to V present the measurements performed for mmWave, circuit switching and software switches. We evaluate the measurements in section VI.

II. Test Setup

Here, we provide a short summary of test setups used in the field and describe the test setups used for the mmWave devices and software switches.

A. Test methodologies

Telecom operators nowadays implement active probe systems for network monitoring and service assurance. By means of these active tests, both time and packet metrics (like latency, jitter, packet loss, reordering or duplication) can be obtained. Additional measurements could include the accuracy of the clock signal. Typical testing measurements are Layer 2 [7] and Layer 3 [8], [9] monitoring, the Two-Way Active Measurement Protocol [10], and validation of throughput capacity and performance of different service classes across the BH, for service activation according to [11], [12], and testing on the actual service configuration. The latter tests validate both configuration and service performance.

B. mmWave devices and test setup

We investigate outdoor and indoor and mmWave scenarios, employing different mmWave devices supporting IEEE 802.11ad 60GHz and connected via USB 3.0 interfaces to the transport node.

Outdoor wireless cards use the Peraso W110 chipsets with multiple carrier modulation and coding schemes (MCS), from MCS0 to MCS12, forward error correction (FEC), and beamforming. The maximum achievable throughput is limited to 4.62Gbps due to the USB3.0 limitation of 5Gbps transfer rate. This chipset uses a SencityMatrix V-band antenna with a high gain of 38dBi and a narrow 3dB beamwidth of 1.8°. Thus, signals radiate with very high directivity and a good transmission range up to 500m.

The indoor wireless chipsets are based on the TC60G USB solution, which comprises a 11mmx15mm module with highly integrated ultra-low power 802.11ad 60Ghz CMOS embedded antennas in 2x2 MIMO configuration. The physical layer (PHY) supports three 802.11ad channels and seven single carrier MCS. The media access control layer (MAC) supports automatic beamforming operation.

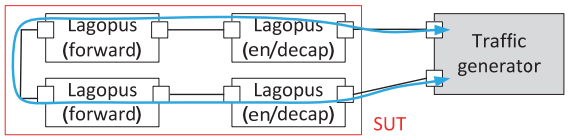

Fig. 2: Test topology for latency/jitter measurements

In the indoor and outdoor test setups, we mount two mmWave devices in two off-the-shelf Linux nodes forming a Line of Sight (LOS) link. We conduct two sets of experiments to characterize the mmWave link performance. One set focusing on the link performance for both indoor and outdoor setups, and another set specifically focusing on the beamforming training capabilities in the indoor link. The outdoor antennas have a fixed beamwidth of 1.8° and are aligned manually.

C. Software switches and test setup

Latency and jitter of OpenFlow (OF) [13] hardware and software switches – here OpenvSwitch [14] – are evaluated in [15]. For the last few years, the Data Plane Development Kit (DPDK) [16] has been widely used as an efficient framework for packet processing in user space. In this paper, we focus on the DPDK based switch Lagopus v0.2.10 [6] to measure latency and jitter. We determine latency and jitter by sending bidirectional traffic with constant frame size and bitrate through the system under test (SUT). There is an additional unidirectional measurement flow, which is timestamped in hardware. The tool moonGen [17] generates the test traffic, controls the hardware timestamping, and collects the latency measurements. We measured minimal and maximal latency of the test traffic to compute packet latency variation or jitter. These tests are considerably simpler to perform than tests following RFC2544 [12], but are sufficient to investigate the latency and jitter metrics.

To measure latency and jitter of the Lagopus software switch we connected one to four instances in a chain to the traffic generator as depicted in Fig. 2. Depending on the test, the two switches connected directly to the traffic generator either forwarded the test traffic or en-/de-capsulated it first.

We tested with both 1G and 10G links. The 10G links were deployed on rack servers with six core CPUs operating at 1.7GHz; the 1G links were deployed on PCs with dual core CPUs and hyperthreading operating at 3.3GHz. The operating system was Ubuntu 14.04 (servers) and 16.04 (PCs), configured to isolate the processor cores for forwarding from other tasks. The network interface cards (NIC) were Intel XL710 for 10G, 82574L for Lagopus and I350 for moonGen as 1G interfaces.

III. mmWave Link Tests

Wireless mmWave is a candidate technology for BH and FH when fiber is unavailable. Under denser 5G deployments, with a need to connect more and more base stations, mmWave links are expected to be used even more often. However, they have to provide the high bandwidth and low latency requirements of 5G

BH and FH traffic. In this section, we provide throughput and jitter measurements of two different mmWave links, one targeting an outdoor and one an indoor scenario.

A. mmWave Transmission (indoor and outdoor scenarios)

Table III.1 summarizes the baseline measurement results of the outdoor mmWave link. The measurements were conducted in an outdoor environment over 185 meters with a fixed beamwidth of 1.8°. The average measured latency of the link is 700µs. We have measured the packet loss rate for both DL/UL user datagram protocol (UDP) and transmission control protocol (TCP) traffic, under different MCS specified in IEEE 802.11ad. In general, TCP traffic exhibits better reliability (in terms of packet loss rate) than UDP at the expense of a lower data rate. This is due to the retransmission and congestion avoidance mechanisms of TCP. Jitter is generally low (around 10µs to 30µs) in all cases.

Table III.1: mmWave outdoor link performance

MCS	DL UDP/TCP Mbps (pkt loss %)	UL UDP/TCP Mbps (pkt loss %)	Jitter DL/UL µs
9	1700/910 (0.87/0)	1340, 1100 (1.1/0)	24/11
8	1220/972 (0.23/0)	1300, 1070 (0.69/0)	20/12
7	1080/894 (0.76/0)	1160, 966 (1.1/0)	29/7
6	910/781 (1.1/0)	986, 856 (0.65/0)	25/11
5	866/760 (0.28/0)	869, 760 (1.1/0)	29/14
4	814/719 (0.34/0)	812, 718 (0.96/0)	13/29
3	705/633 (0.26/0)	709, 633 (0.39/0)	20/20

Table III.2 reports the baseline measurement results of the indoor mmWave link. The numbers indicate the maximum achieved UDP throughput and distance for the 7 MCS available in TC60G-USB3.0 under perfect antenna alignment conditions.

Table III.2: mmWave indoor link performance

MCS	Achieved MPDU Throughput (Mbps)	Achieved Phy Rate (Mbps)	Max Distance (meters)
7	1020	1925	2.6
6	900	1540	3.6
5	790	1251.25	4.8
4	760	1155	8.2
3	665	962.5	10.3
2	565	770	15.7
1	325	385	16.9

Packets of 4096 bytes are transmitted for all 7 MCS. The results show a clear throughput drop with distance. At short distances the MAC protocol data unit (MPDU) throughput is close (around 1Gbps) to the one expected with MCS7 (QPSK modulation). For the longest distances covered by the 10dBi antennas (around 17-18 meters) MPDU throughput decreases (around 325Mbps) like the one expected (around 320Mbps) with MCS1 (BPSK modulation). The results also indicate a significant bias between the MPDU throughput and PHY Rate. This bias is due to the excessive retransmissions needed to decode at the receiver side with maximum distance.

Note that higher distances can be achieved by using higher gain antennas as proved in the outdoor scenario. The embedded 10dBi antennas emit an Equivalent Isotropically Radiated Power (EIRP) of just 14dBm due to the limited power consumed by this ultra-low power equipment. The low power consumption of these antennas fed through an USB 3.0 port explains the limited indoor distances covered by these devices.

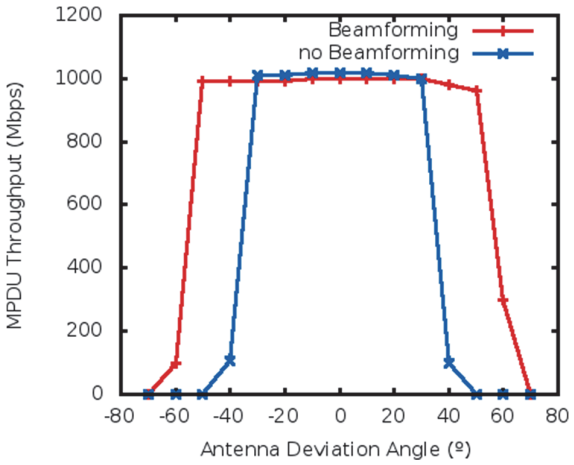

Fig. 3: mmWave indoor beamforming perf., MCS 7

The measured throughput indicates that depending on the throughput requirements of the requested functional split and the actual traffic load, higher layer applications can consider to change the MCS to reduce transmission power.

B. Impact of beamforming mode on mmWave links

The indoor mmWave device described in Section II.B supports automatic beamforming procedures. Fig. 3 shows the achieved UDP throughput with and without beamforming procedures enabled and MCS 7, which is the MCS most sensitive to antenna misalignment effects over a distance of 1.2 meters. For each experiment, we inject packets of 4096 bytes from the transmitter to the receiver. One of the mmWave antennas was placed with a variable deviation angle with respect to the direction pointed by the other antenna, indicated by the x-axis in Fig. 3. As expected, the results confirm the convenience of enabling beamforming training procedures. We observe that although peak MPDU throughput rates are achieved with a small deviation angle and no beamforming enabled, the overhead required by the beamforming training procedures causes a throughput reduction of merely 2% (around 1020Mbps versus 1000Mbps). On the other hand, the results show excessive throughput inefficiencies (990Mbps against 105Mbps for 40º) without enabling beamforming procedures.

IV. CIRCUIT & OPTICAL SWITCHING

In the 5G-Crosshaul data plane architecture, see Fig. 1, the forwarding elements are multi-layer elements, which may contain both a packet and a circuit forwarding element. The circuit part is itself composed of a Time Division Multiplex (TDM) layer on top of an optical switch (for these measurements we used a Dense Wavelength Division Multiplexing, DWDM, optical link), as described in [5]. This new approach allows the operator to transparently move packets between the packet and optical domain, depending on the requirements of the flow, especially on latency, and to control this allocation of flows to the packet or circuit domain following an SDN approach. Overall, the infrastructure would be used more efficiently, resulting in energy-savings. To evaluate the solution, we measured the additional latency introduced by the multiplexing on CPRI flows (the FH source with the most stringent requirements), and the Bit Error Rate (BER).

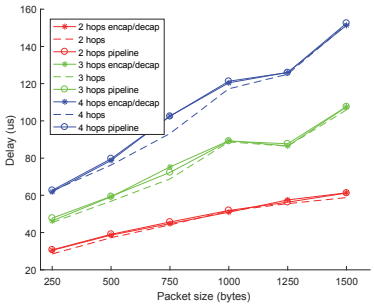

Fig. 4: Avg latency, rate 250Mbps

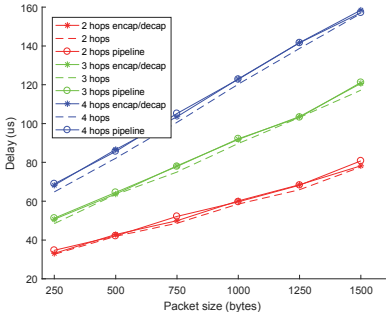

Fig. 5: Avg latency, rate 500Mbps

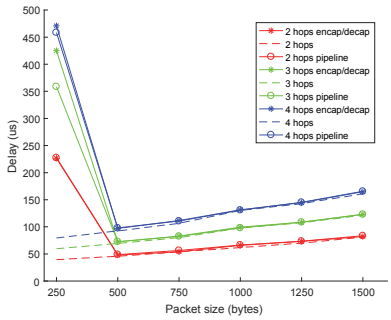

Fig. 6: Avg latency, rate 750Mbps

In the test setup for latency, a CPRI BBU is connected to a RRH by a XCSE operating over a DWDM link of 14Km length. We connect two CPRI pattern analyzers to the ingress CPRI link on one side and the egress CPRI link on the other side. These analyzers detect the start of the hyperframes within the CPRI frames. The time elapsed between the hyperframe start on the two sides minus the fiber propagation delay (determined by fiber length, type and wavelength) is the additional latency introduced. We measured an end-to-end latency introduced by the XCSE nodes, including FEC, of 4µs, and a BER after FEC of less than $< 10^{-15}$.

V. SOFTSWITCH MEASUREMENTS

In general, software switches are more flexible than hardware switches. Due to recent performance improvements, software switches can be used instead of hardware switches following current network virtualization trends. We performed a series of tests to evaluate the impact of different configurations of the underlying packet processing, of the flow tables, and the number of hops.

A. DPDK burst size

The software switch Lagopus is based on DPDK [16], which avoids the overhead of interrupts by actively polling the NIC. Different DPDK threads exchange frame descriptors through ring buffers. To reduce the overhead of accessing the NICs and the ring buffers, DPDK and Lagopus enqueue or dequeue a burst of frames at once to or from the ring buffers. For each of the three main tasks – receiving, processing and transmitting – the burst sizes of reading from the NIC or ring buffer and of writing to the NIC or ring buffer can be configured.

Configuring large burst sizes increases throughput as the cost of accessing data structures is shared among several packets. But processing frames in bursts increases latency and especially jitter (default burst size is 32). To reduce latency and jitter we have configured all burst sizes for writing to the NIC or to ring buffers to 1, thus, frames are forwarded as quickly as possible. To allow some amount of sharing costs among packets we configured all read burst sizes to 8 or 16. Measurements using 2.5Gbps of traffic on each direction for frame sizes of 512Bytes result on an average latency (with write burst sizes of 1) of about 5.4µs compared to 7.2µs with write burst sizes of 2.

Despite the small write burst size, a throughput of more than 7.5Gbps was achieved on the rack servers. This throughput would allow to aggregate packetized CPRI traffic of three RRHs on one 10G link.

For comparison, we also measured latency and jitter for two directly connected ports of the traffic generator and for the DPDK *testpmd* application, which we consider the simplest DPDK based application to exchange frames between two ports. The average latency between two directly connected 10G ports was 0.35µs with a maximum of 0.37µs. The measured end-to-end latency through the *testpmd* application on 10G ports, configured with burst size 4, was 4.3µs and a standard deviation of 2.6µs. The large standard deviation shows the impact of processing packets in bursts.

We observed a large variation of jitter, from 20µs to 200µs, on different PC platforms. Especially the available rack servers caused significantly higher jitter than standard PCs, despite using similar operating systems or processors. This increases the configuration effort of the hosts or for tuning of operating system kernels. This issue is relevant as well for virtual network function with real-time or close to real-time requirements.

B. Pipeline depth

The pipeline of an OF switch may consist of more than a single flow table. Multiple flow tables may be used to avoid specifying all combinations of header fields in one flow table and to reduce the size of flow tables and the effort of modifying them. But using a sequence of flow tables requires additional computational resources and as such increases the latency.

We extended the exchange of frames among ports by a chain of flow tables, flow table n just containing a single action *goto-table*(*n+1*), without any match field. The effort for such a *goto-table* action is the minimum effort to use an additional table. The average latency increased from about 5.4µs for 1 *goto-table* instruction to about 6.6µs for 20 *goto-table* instructions. The difference between 1 *goto-table* and 20 *goto-table* actions is 1.2µs, each *goto-table* action contributing about 60ns latency.

C. Multiple hops

Previous measurements focused on the operation of a single switch. Here, we measure the performance of the 5G-Crosshaul XCF encapsulation and the processing delay of the 5G-Crosshaul OF pipeline in a testbed of 4 switches in a daisy chain topology. Each switch was deployed on a PC with 4 logical cores. Two of the logical cores are used for forwarding and accessing the NICs, the other two cores are used for general Linux tasks and control of the switch.

We used three flow table setups to compare the effects of the 5G-Crosshaul innovations on switch performance regarding three metrics: latency, packet loss rate, and jitter. The first flow table setup is OF forwarding among pairs of ports. The second one is OF forwarding among pairs of ports with hardcoded XCF en-/decapsulation in a single rule ('encap/decap' in Fig. 4 to Fig. 6). The third one is the 5G-Crosshaul OF pipeline ('pipeline' in Fig. 4 to Fig. 6) for forwarding with enhanced en-/decapsulation and multi-tenant support.

Fig. 4 to Fig. 6 show the latency measurements. As expected, the delay increases with the number of hops and the packet size. The increase with the size of packets is due to the serialization delay on 1G links. As observed in the figures, the XCF e/d and the Pipeline processing is not affecting significantly the latency of the packets. Fig. 6 shows a delay increase for small packets, corresponding to a large packet rate. The resulting packet rate saturated already the switch with a corresponding increase of the average latency. These measurements and the analysis of the delay distribution are described in detail in [18].

D. Flow table modification

The switch accesses the flow table both when forwarding frames and when changing flow entries by an SDN controller. This might block forwarding temporarily and thereby increase jitter. To measure the impact of flow table modifications we simulated a moderate load of changing flow tables by adding and deleting 6 flows per second. No significant difference in terms of latency and its standard deviation were observed, both values differed by less than 1%.

VI. CONCLUSIONS

This paper presented a first suitability analysis for the transport of BH and FH traffic of the different elements used in the 5G-Crosshaul architecture. We have measured several performance metrics for mmWave links, XCSE (circuit switching over optical) elements and software switches.

In particular, the preliminary measurement results provided in this paper indicate that mmWave links with a narrow beam width of 1.8° are able to support Gbps transmissions over hundreds of meters. The narrow beam nature of mmWave link enables effective spatial reuse and interference avoidance, which is particularly useful in dense urban scenarios where multiple Gbps links need to be concurrently provisioned to cope with increasing traffic demands. For indoor links, the distance can be increased with higher-gain antennas. The beamforming training procedures allow relaxing direction alignment among antennas, easing their deployment. Although throughputs of 1Gbps have been achieved, these are not sufficient for CPRI-like traffic. Also, the reported latency measurements (around 700µs) are still too high for CPRI-like traffic in a C-RAN scenario or for BH traffic of low-latency 5G scenarios. On the positive side, the jitter is already low and bandwidth is high enough to for RAN functional splits at higher layers with lower bandwidth and less stringent delay and synchronization requirements than those posed by CPRI.

As expected, the XCSE or circuit switching part showed a low, constant latency and a low per hop added latency. This demonstrates that this technology is a good candidate to deal with CPRI-like traffic sources, which cannot be dealt with the packet domain due to its stringent requirements.

For software switches, latency and jitter values are low enough to be usable for FH traffic for a small number of hops, especially for functional splits with less stringent requirements than CPRI. This would allow to use software switches close to virtual machines, whereas the field-deployed switches could still be hardware ones. The way of configuring flow tables has a measurable impact on latency. The flexibility of SDN controllers to configure flows should be used to decrease the forwarding delay within switches. Thus, efforts on implementing software switches will continue to reduce jitter values and at the same time keep or even increase throughput.

More work on performance tuning is required to transport FH traffic with CPRI characteristics (one way delay below 100usec, jitter in the order of ns and capacity exceeding 1Gbps). These requirements will be relaxed with other functional split options. Although the exact requirements are not yet defined, we expect that the current performance of 5G-Crosshaul elements is already suitable to be used, even for several hops.

REFERENCES

[1] CPRI Specification V7.0, http://www.cpri.info/spec.html

[2] A. de la Oliva et al., "An Overview of the CPRI specification and its application to C-RAN based LTE scenarios", IEEE Communications Magazine, Febr. 2016

[3] D. Wübben et al., "Benefits and Impact of Cloud Computing on 5G Signal Processing", IEEE Signal Processing Magazine, Nov. 2014

[4] A. D. La Oliva et al., "Xhaul: toward an integrated fronthaul/backhaul architecture in 5G networks", IEEE Wireless Communications, Oct. 2015

[5] F. Cavaliere, et.al., "Towards a unified fronthaul-backhaul data plane for 5G -The 5G-Crosshaul project approach", Computer Standards & Interfaces, Volume 51, March 2017, Pages 56–62

[6] Lagopus, http://www.lagopus.org

[7] ITU-T Recommendation Y.1731, "Operation, administration and maintenance (OAM) functions and mechanisms for Ethernet-based networks", Aug. 2015.

[8] ITU-T Recommendation Y.1540, "IP packet transfer and availability performance parameters", July 2016.

[9] J. Mahdavi, V. Paxson, "IPPM metrics for measuring connectivity", RFC 2678, Sept. 1999.

[10] K. Hedayat, R. Krzanowski, A. Morton, K. Yum, J. Babiarz, "A Two-Way Active Measurement Protocol (TWAMP)", RFC 5357, Oct. 2008.

[11] ITU-T Recommendation Y.1564, "Ethernet service activation test methodology", Feb. 2016.

[12] S. Bradner, J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, March 1999.

[13] OpenFlow, https://www.opennetworking.org

[14] Open Vswitch, http://openvswitch.org/

[15] F. Dürr and T. Kohler, "Comparing the Forwarding Latency of OpenFlow Hardware and Software Switches", Institute of Parallel and Distributed Systems (IPVS), Universität Stuttgart, TR 2014/04, 2014

[16] DPDK, http://dpdk.org/

[17] MoonGen, http://scholzd.github.io/MoonGen/index.html

[18] N. Molner, S. González, T. Deiß, A. D. La Oliva, The 5G-Crosshaul Packet Forwarding Element pipeline: measurements and analysis, submitted to CLEEN17