First Demonstration of Autonomous TSN-based Beyond-Best-Effort Networking for 5G NR Fronthauls and 1,000+ Massive IoT Traffic

Naotaka Shibata(1), Paikun Zhu(2), Kazuto Nishimura(3), Yuki Yoshida(4), Kazunori Hayashi(5), Masaki Hirota(3), Rintaro Harada(1), Kazuaki Honda(1), Shin Kaneko(1), Jun Terada(1), and Ken-ichi Kitayama(2,4)

(1) NTT Access Network Service Systems Laboratories, naotaka.shibata.vr@hco.ntt.co.jp
(2) Graduate School for the Creation of New Photonics Industries
(3) Photonics System Business Unit, Fujitsu Limited
(4) National Institute of Information and Communications Technology
(5) Graduate School of Engineering, Osaka City University

Abstract We propose and demonstrate a beyond-5G x-haul platform with autonomous time aware shaper and adaptive mobile fronthaul compression, which simultaneously accommodates real-time 5G NR-class fronthauls (2*29.49-Gbps CPRI-equivalent-rate) and 1,000+ massive IoT live backhaul traffic through a 10GbE interface while keeping <100-μs end-to-end fronthaul latency.

Introduction

5G cellular services have recently been launched globally. The initial service based on 3GPP Release 15 mainly focused on the enhanced mobile broadband (eMBB)[1]. The evolved service, beyond 5G, additionally supports ultra-reliable and low latency communications (URLLC) and massive machine type communication (mMTC)[2]. Such use cases are required by verticals and industrial Internet of Things (IoT). Many of them expect to construct non-public networks[3] for exploiting dedicated and secured networks.

For small and medium enterprises, a 5G public network integrated non-public network[4] is a cost-effective solution. However, there are various requirements from x-haul[5], and they need to be satisfied in the network at the same time. Techniques to assure a low latency have been discussed in time sensitive networking (TSN)[6-8] and beyond best-effort network[9]. On the contrary, the accommodation of massive devices that causes a significant fluctuation of required bandwidth should be addressed. In this context, the network must be flexible for the addition and change of various requirements, and enormous configurations by the network operator accompanying them should be eliminated by network automation.

In this paper, we propose and demonstrate a beyond-5G x-haul platform with autonomous time aware shaper (TAS) and adaptive mobile fronthaul (MFH) compression, which simultaneously accommodates 5G NR-class MFHs (2*29.49-Gbps CPRI-equivalent-rate) and 1,000+ massive IoT live backhaul traffic through a 10 GbE interface while keeping <100-μs end-to-end MFH latency. It is noted that under COVID-19, we carried out the joint experiment remotely by connecting the network elements through a nation-wide R&D network testbed, namely JGN[10].

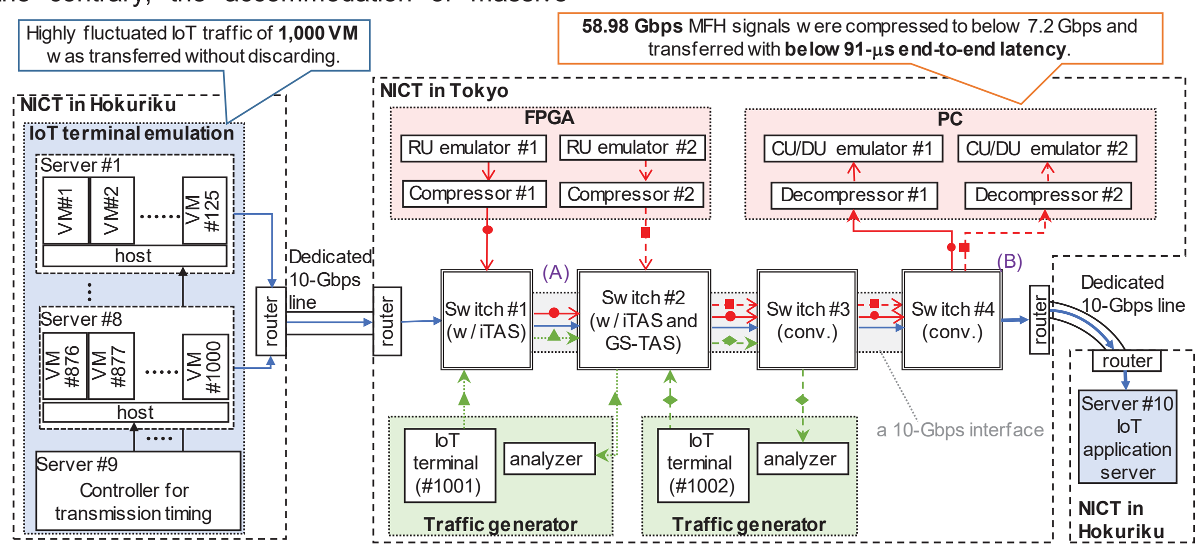


Fig.1: Overview of testbed and our achievement.

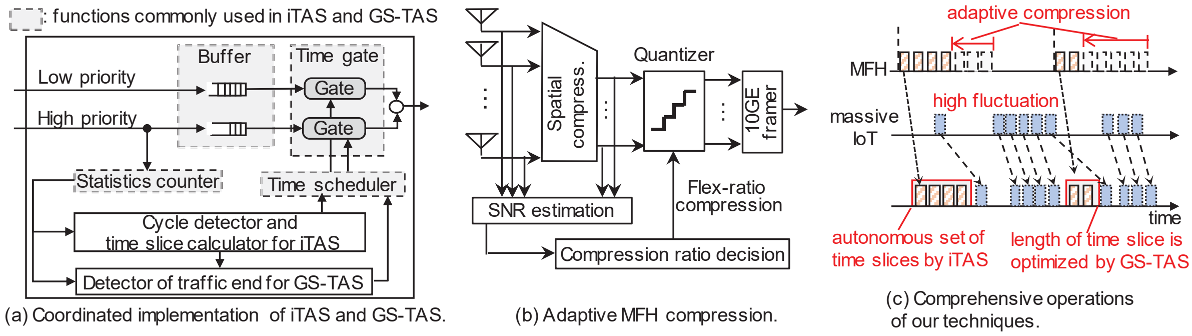


Fig.2: Block diagrams of key enablers and comprehensive operations.

Testbed and Key Enablers

Figure 1 shows the overview of the testbed and our achievements. The dedicated 273-km 10-Gbps line, JGN, is used to connect Tokyo and Hokuriku. Based on conventional TAS technique[11], we propose an autonomous operation of it, and implement the proposal with Switches (SWs) #1 and #2. The key advantage of the TAS is that it is required only at the aggregation point of high-priority and low-priority traffic. Therefore, SWs #3 and #4 are conventional ones without TAS to demonstrate the integration of SWs with TAS. Each SW is connected with only a 10-GbE interface. We emulated remote units (RUs) by FPGA and generated two 29.49-Gbps MFH signals having the short transmission time interval (TTI) of 250 μs in 5G NR. Key enablers for the proposed system are the followings.

IoT signal emulation: We prepared 10 physical servers at NICT in Hokuriku[12]. 1,000 virtual machines (VMs) are launched using 8 servers. Each VM sends the data to an IoT application server (server #10) periodically assuming sensor applications. To cause the largest possible fluctuation in IoT traffic, the transmission timing of VMs was controlled by a control server (server #9), so that all VMs simultaneously transmit the traffic. In Tokyo, the traffic generator outputs two 1-Gbps traffic as IoT terminals (#1001 and #1002) assuming application of real-time surveillance.

iTAS: We propose an intelligent TAS (iTAS) in order to autonomously set the time slices for high-priority, that is, latency-sensitive traffic. In each time slice, only the latency-sensitive traffic is transferred while non-latency-sensitive traffic cannot be transferred until the time slice ends. The cycle of the traffic is detected by calculating autocorrelation at the SWs.

Although, the conventional TAS requires time synchronization among SWs, iTAS does not require time synchronization schemes at all. The iTAS can also determine the timing of the time slice without considering the processing delay as well as fiber propagation delay. Hereafter, MFH and IoT are introduced for latency-sensitive and non-latency-sensitive traffic, respectively. In the case of MFH, the traffic is generated periodically by the wireless scheduling.

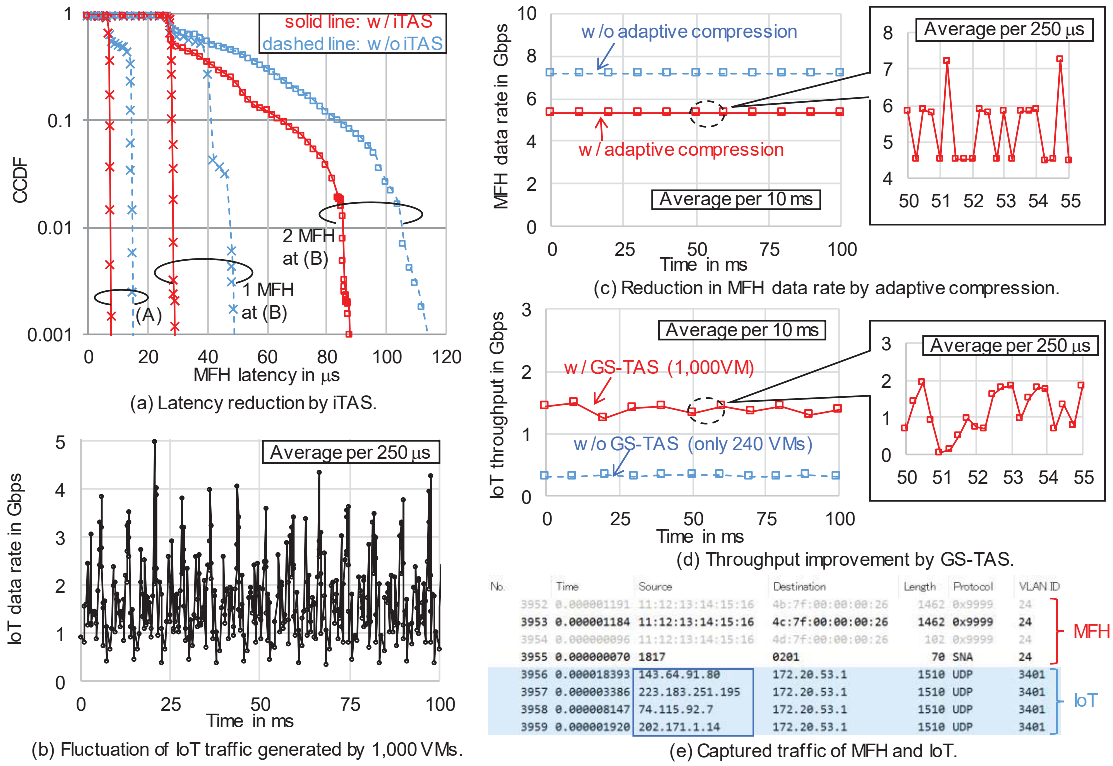
GS-TAS: The length of the time slice calculated by iTAS is sufficiently long to include the variance in the amount of traffic. Our proposed gate shrunk TAS (GS-TAS) optimizes the length of the time slice automatically every cycle[13]. By the GS-TAS, the IoT traffic starts to be transferred immediately after the MFH traffic ends, and thus the throughput of the IoT traffic improves. In this experiment, we implemented the coordinated function of iTAS and GS-TAS in SWs by using FPGA as shown in Fig. 2(a).

Adaptive MFH Compression: In order to significantly relax the MFH bandwidth requirement with low latency, we have proposed adaptive space-time analog-to-digital-compression (ADX)[14], where MFH data can be reduced by ~90% while still supporting up to 1024QAM wireless modulation and maintaining <500 μs processing latency. In this experiment, compressed MFH data are encapsulated to IEEE802.3-compatible packets for TSN networking. Moreover, adaptive compression ratio facilitates GS-TAS, which can be provisioned by central unit/distributed unit (CU/DU) or even autonomously, such as by signal-to-noise ratio (SNR) sensing via subspace tracking as shown in Fig. 2(b).

Figure 2(c) shows the coordinated operations by key enablers. The iTAS autonomously assures a low latency for the compressed MFH signals. The GS-TAS optimizes the length of each time slice according to the time-varying MFH traffic by adaptive compression. By using its enhanced bandwidth, massive IoT traffic is transferred by exploiting statistical multiplexing gain without any bandwidth allocation.

Results

The complementary cumulative distribution function (CCDF) of MFH latency is shown in Fig. 3(a). The CCDF shows the probability of the MFH



No.	Time	Source	Destination	Length	Protocol	VLAN ID
3952	0.000001191	11:12:13:14:15:16	4b:7f:00:00:00:26	1462	0x9999	24
3953	0.000001184	11:12:13:14:15:16	4c:7f:00:00:00:26	1462	0x9999	24
3954	0.000000096	11:12:13:14:15:16	4d:7f:00:00:00:26	102	0x9999	24
3955	0.000000070	1817	0201	70	SNA	24
3956	0.000018393	143.64.91.80	172.20.53.1	1510	UDP	3401
3957	0.000003386	223.183.251.195	172.20.53.1	1510	UDP	3401
3958	0.000008147	74.115.92.7	172.20.53.1	1510	UDP	3401
3959	0.000001920	202.171.1.14	172.20.53.1	1510	UDP	3401

(e) Captured traffic of MFH and IoT.

Fig. 3: Experimental results.

latency exceeding a given latency. At (A) in Fig.1 (between SWs #1 and #2), the iTAS reduced the latency and assured a deterministic latency.

At (B) in Fig. 1 (at the output of SW #4), by temporally stopping the RU emulator #2, the latency of only 1 MFH signal was evaluated at first. Although MFH signals passed through conventional SWs, the deterministic latency was obtained. Then, by activating the RU emulator #2, two MFH signals with the same priority were aggregated, resulting in the latency fluctuation. Even so, the iTAS reduced the latency. That is to say, the iTAS autonomously decided the cycle and length of time slices according to MFH traffic. The maximum end-to-end latency at (B) in Fig. 1 with the iTAS was 91 µs which meets the requirement of 100 µs[6], while the latency without iTAS was 117 µs.

Figure 3(b) shows the data rate of IoT traffic generated by 1,000 VMs in Hokuriku. While the average data rate was 1.4 Gbps, the maximum data rate was 5.0 Gbps; the highly fluctuated traffic was generated.

Figure 3(c) shows the total data rate of 2 MFH traffic. The ADX compression ratio of each MFH was adaptively changed between 8% and 13% per 250-µs TTI according to the instantaneous SNR of wireless signals. Enabled by the adaptive ADX, the average MFH data rate was successfully reduced compared with conventional fixed-ratio compression. We achieved 1.49 % average error vector magnitude (EVM) for 16-channel MIMO 1024QAM MFH signals after TSN transport. It include a slight 1.08 % EVM penalty by our compression.

Figure 3(d) shows the IoT throughput at (B) in Fig. 1. When the GS-TAS was off at SW #2, only 240 VMs could be accommodated without loss of frames. By using GS-TAS, IoT throughput was improved since the length of each time slice was optimized according to the time-varying MFH traffic by adaptive compression. In this case, we confirmed that the traffic of 1,000 VMs was transferred without loss of frames. At the same time, MFH and IoT #1001–#1002 were simultaneously transferred with no frame loss. Figure 3(e) shows the screen shot of the traffic analyzer around the end of time slice. This figure shows that various VMs with different source IPs were transferred after the end of MFH traffic .

Conclusions

Autonomous TAS-based networking has been demonstrated for 58.98-Gbps MFH signals with 91-µs end-to-end latency and over 1,000 IoT terminals by coordination of multiple techniques. This envisages coordinated networking of public and non-public traffic for beyond 5G.

Acknowledgements

A part of this work was conducted under the R&D contract, "Wired-and-Wireless Converged Radio Access Network for Massive IoT Traffic (JPJ000254)" funded by the Ministry of Internal Affairs and Communications, Japan. We thank NICT for the use of StarBED, JGN, and the experimental supports.

References

[1] A. Ghosh, A. Maeder, M. Baker, and D. Chandramouli, ‘‘5G evolution: A view on 5G cellular technology beyond 3GPP release 15,’’ IEEE Access, vol. 7, pp. 127639–127651, 2019.

[2] NTT DOCOMO: “5G Evolution and 6G White Paper,’ Jan. 2020.

[3] “5G Non-Public Networks for Industrial Scenarios,” March 2019 [online] Available: https://www.5g-acia.org/index.php?id=6958.

[4] 3GPP, TS 23.501, v16.6.0 Sep. 2020.

[5] Y. Yoshida, “Mobile xHaul evolution: enabling tools for a flexible 5G xHaul network,” OFC2018, Tu2K.1 (Tutorial).

[6] Time-Sensitive Networking for Fronthaul, IEEE Standard 802.1CM Draft 2.2 Mar. 2018.

[7] N. Benzaoui, M. S. Gonzalez, M. V. Rivera, J. M. Estarán, H. Mardoyan, W. Lautenschlaeger, U. Gebhard, L. Dembeck, Y. Pointurier, and S. Bigo, “DDN: Deterministic Dynamic Networks,” ECOC2018.

[8] J. Zou, et al., “Options for time-sensitive networking for 5G fronthaul,” ECOC, M.1.F.5, Sep. 2019.

[9] A. Clemm, M. F. Zhani, and R. Boutaba, “Network management 2030: Operations and control of network 2030 services,” Journal of Network and Systems Management, pp. 1–30, 2020.

[10] https://testbed.nict.go.jp/jgn/english/index.html

[11] Time Aware Shaper, IEEE Standard 802.1Qbv (Draft 3.1) Oct. 2015.

[12] https://starbed.nict.go.jp/en/index.html

[13] N. Shibata, S. Kaneko, K. Honda, and J. Terada, “Deterministic Layer-2 Ring Network with Autonomous Dynamic Gate Shaping for Multi-Service Convergence in 5G and Beyond,” OFC2020, Th2A.36.

[14] P. Zhu, Y. Yoshida, and K. Kitayama, J. Lightwave Technol., DOI 10.1109/JLT.2020.3029271.