Session Management across Heterogeneous Wireless Technologies in a Rail Transport Environment

Panagiotis Karamichailidis*, Nikos Makris*, Paris Flegkas*, Anna Tzanakaki†, Markos Anastasopoulos†, Ioanna Mesogiti§, George Lyberopoulos§, Christina (Tanya) Politi‡, Christos Tranoris‡, Thanasis Korakis*
*University of Thessaly, Volos, Greece
†National and Kapodistrian University of Athens, Athens, Greece
§COSMOTE Mobile Telecommunications S.A., Athens, Greece
‡University of Patras, Patras, Greece

***Abstract*—5G networks are transforming everyday's living through the introduction of disruptive technologies in the network, enabling a rich ecosystem of services. One of the vertical sectors that can benefit from the latest activities around 5G communications is the railway transport industry, for providing enhanced services to on-board passengers and supporting time-critical applications for operations management. The H2020 5G-VICTORI project is working on delivering a fully-fledged 5G solution, directly applicable to railway environments, using heterogeneous wireless high-capacity links connecting trains to the network. In this work, we focus on the handover management across multiple heterogeneous networks, that are deployed on the track-side and facilitate track-to-train communications. We use P4 programming to configure the network switches, with the aim to present the applications that run on the train with a seamless handover experience, regardless of the dynamics of each heterogeneous link. The paper describes the P4-based session management process, and presents a first feasibility evaluation of the framework in a testbed environment using software switches.**

***Index Terms*—5G, Railway Communications, P4, mobility management**

I. Introduction

5G networks and their evolution are transforming the telecommunication infrastructure landscape through their inherent support of enhanced Mobile Broadband (eMBB), ultra-reliable low-latency (uRLLC) and massive Machine Type Communications (mMTC). The support of such services drives the transformation of several sectors vertical to the telecommunications industry, through new services/applications requiring ubiquitous high-speed connectivity, reliability and dense deployments. An excellent example of such a vertical sector which stresses the network requirements is the railway industry, aspiring to support a number of on-board services, assisted through track-side deployments.

Contributions around Connected and Automated Mobility are key for the success of the Railway sector. In particular, efforts focus on the Future Railway Mobile Communication System (FRMCS), as a system that drives the digitalization of the rail transport industry, which can benefit from the adoption of several 5G concepts and technologies. Through track-side deployments, high-bandwidth connections can be established with moving trains, delivering multi-Gbps speeds per wagon. Such connections can be leveraged for a wide set of services, namely Business (e.g. infotainment), Performance and Critical services [1]. Critical services include the transmission of signaling needed for coordinating the access of trains across crossings, or security by means of track monitoring systems. As such, stringent KPIs need to be met for the smooth operation of services on top, through separate slices for low latency communication and high-bandwidth. FRMCS requirement for technology neutrality at various layers is driving the heterogeneous network deployments (at the transport and access layers). To this end, session continuity over different track-to-train technologies is needed for achieving uninterrupted service provisioning.

Given the aforementioned challenges, the 5G-VICTORI project [2] is designing and implementing such a solution for demonstration in an operation environment. The solution builds upon the demonstration of the 5G-PICTURE project [3] in a railway environment, extending it with multi-technology track-to-train communications, and a novel mobility management framework that is applicable to high-speed heterogeneous wireless track-side networks with varying characteristics. The project is working on delivering an end-to-end solution for uninterrupted communications of the train with the edge and cloud deployments, on top of which critical and business services will be executed. This paper is describing the network design and provides a first experimental evaluation of a session continuity solution between the heterogeneous solutions that will realize the track-to-train communication.

The rest of the paper is organized as follows: Section II is describing the end-to-end network deployment, and provides some information on the vertical services that will be provided on top. Section III details the solution that will realize session continuity across different access points, relying on network programmability. Section IV shows a first evaluation on the feasibility of the scheme, and some experimental results. Finally, in Section V we conclude the paper.

Corresponding authors: Panagiotis Karmichailidis (karamiha@uth.gr), Nikos Makris (nimakris@uth.gr), Paris Flegkas (pflegkas@uth.gr)
The research leading to these results has received funding by the European Horizon 2020 Programme for research, technological development and demonstration under Grant Agreement Number 857201 (H2020 5G-VICTORI). The European Union and its agencies are not liable or otherwise responsible for the contents of this document; its content reflects the view of its authors only.

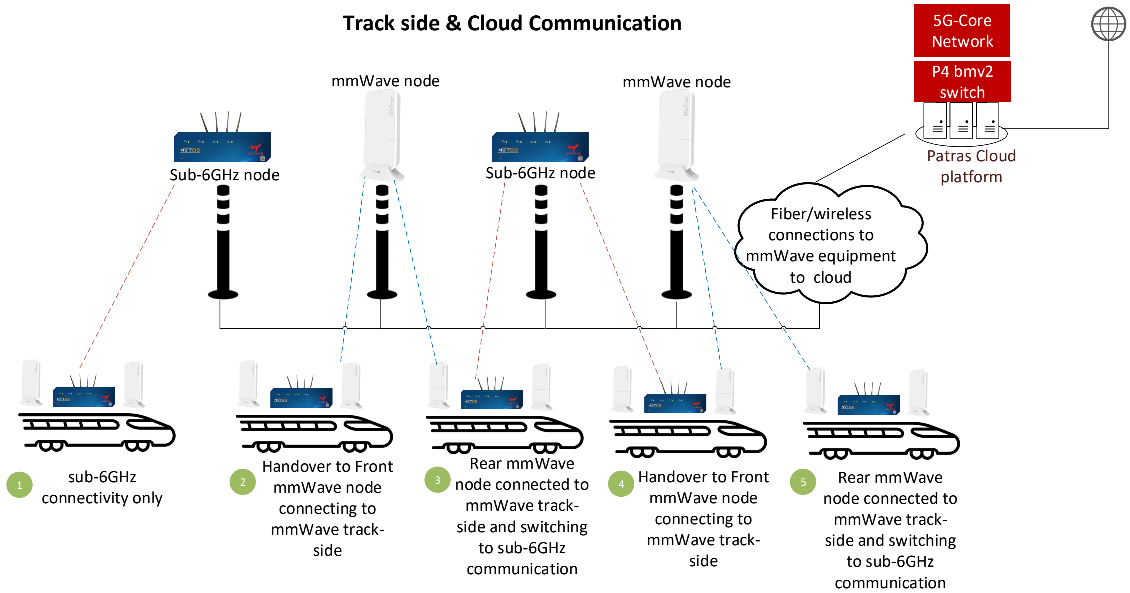

Fig. 1. Track-to-train and track-to-cloud communication network for the 5G-VICTORI railway setup; sub-6GHz are realized using IEEE 802.11ax nodes, and mmWave using 60GHz wireless units.

II. End-to-End Network Deployment

The 5G-VICTORI train deployment in the city of Patras builds upon the functionality of the 5G-VINNI testbed [4], appropriately extended to support the use case by including track-side and on-train deployments of equipment. The setup will address wireless connectivity for on-board, track-to-train and track-to-cloud. Figure 1 shows the track-side deployments for providing connectivity to a moving train, and how the connection to the cloud is established. Heterogeneous Access Points (APs) will be deployed in pairs for the track-to-train communication, addressing sub-6GHz (using IEEE 802.11ax solutions) and mmWave (60GHz) frequencies. For the sub-6GHz nodes, omni-directional antennas will be used, for covering a wider sector, whereas for the mmWave APs, directional 60-degree sectors are used. All the track-side APs are interconnected by either fibre, or other mmWave equipment that provides their connection to the cloud. The proposed solution is progressing beyond other proposed deployments such as in [5], by addressing access gaps in coverage that might exist due to the high directionality of mmWave links taking advantage of heterogeneous networks, while discovery of the units that the train is associated with is achieved without packet duplication, as detailed below.

Towards providing connectivity on-board, three APs will be deployed on the rooftop of the train. Given the high directionality of the mmWave APs, two of them will be deployed on the train roof, one in the front and one in the rear part of the wagon, towards ensuring that at least one of them is always under the coverage of a track-side unit. As sub-6GHz will use an omni-directional setup, only one node suffices to connect to the respective track-side units.

On-board, the deployed equipment reflects the use cases that will be supported. These include the following:

- **Track monitoring with on board cameras video streaming to Control center over 5G Rail network deployment**: track will be constantly monitored via high-resolution cameras, streaming their footage in the cloud, where analysis of the footage will take place in real-time.
- **Mission-Critical Push to Talk (MCPTT) for railway operations**: MCPTT equipment will be integrated in the overall setup, enabling communication among the railway staff.
- **5G data services for passengers**: a disaggregated 5G-cell and a WiFi AP will be deployed on board, for providing high-speed data services to connected clients. The moving cell will be backhauled to the 5G core (located at the cloud platform) over the network, and will use the OpenAirInterface platform [6].

Figure 2 details the on-board network and equipment used. In such a dynamic environment, train mobility poses several challenges for the handover process between the heterogeneous track-side APs, as the different technologies support different rates for transmission of data, and different delay characteristics. Figure 1 shows the possible handover decisions that need to take place between the train and the track-side units. Towards overcoming these challenges, Software Defined Networking (SDN) is employed, leveraging the features of high programmability that it offers. Since flow establishment time can affect the handover process and cause drops in the connection, we employ P4-switching [7] that enables high speed configuration and wide programmability of the networking fabric, contrary to legacy solutions such as OpenFlow. For an initial deployment, we use the P4 behavioral model software switch (bmv2) [8], deployed at two different places: on-board, and at the cloud before accessing the services and/or the Internet.

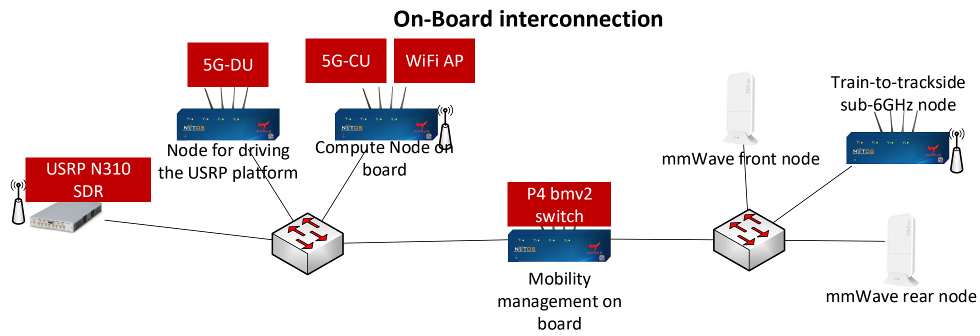

Fig. 2. On Board network setup enabling the execution of vertical services

III. Mobility Management for Session Continuity

SDN switching is deployed at two places: on-board and the cloud platform. The two different points are needed for changing the Ethernet headers per packet, for providing seamless end-to-end connections, regardless of the handover processes. P4 switching has been opted for the use case as it enables a wide variety of actions for switching and can address high workloads of traffic, in the order of multiple-Gbps [9].

The cloud-located controller is used to address the track-side unit that will be used for the downlink channel. Addressing the track-side unit is based on a separate VLAN interface, that each of the units has assigned, which the switch will push/pop on every packet. Selecting the appropriate track-side unit used to forward the data relies on determining which unit has an established connection with the train at each time-point. In order to determine this, the Simple Network Management Protocol (SNMP) is used; each AP can be queried with an SNMP Object Identifier (OID) for retrieving the connected clients. In such a case, the handover time is equal to the time needed to retrieve the information from the APs (in the order of msecs), and the time needed to establish the flows (in the order of usecs). Although the mmWave equipment supports this functionality off-the-shelf, for the sub-6GHz equipment there is no such support for off-the-shelf open source solutions. Several commercial platforms exist that provide such functionality, however, the served OIDs are not providing suitable information for managing the handover process. Therefore, we have implemented such support for the open-source platforms which are used by the project (802.11ax chipsets compatible with the *ath11k* driver) by extending the Management Information Base (MIB) with a respective OID that holds the information of the connected clients. Such a scheme relies heavily on the time-point that the server is queried from the client side in order to respond with the attached clients. Since this granularity can widely affect the handover performance, the sub-6GHz units are extended with callback processes through the SNMP trap feature; whenever there is a change in the associated clients, the trap interface on the SNMP agent sends the message to a set of designated clients (e.g. the network controller on the software switch) without waiting to be queried. This feature drastically reduces the time that the network controller realizes that a handover process has happened.

A similar process is also employed for the uplink case, on the software switch on-board. Using the SNMP messages that address the on-board nodes, we can determine which links are active and their quality of connection. The on-board controller subsequently selects the device (sub-6GHz or front/rear mmWave) that will be used for transmitting data to the core. The controller can be further extended with slicing support for the hosted services, e.g. by prioritizing data from more critical communications like MCPTT over the backhaul.

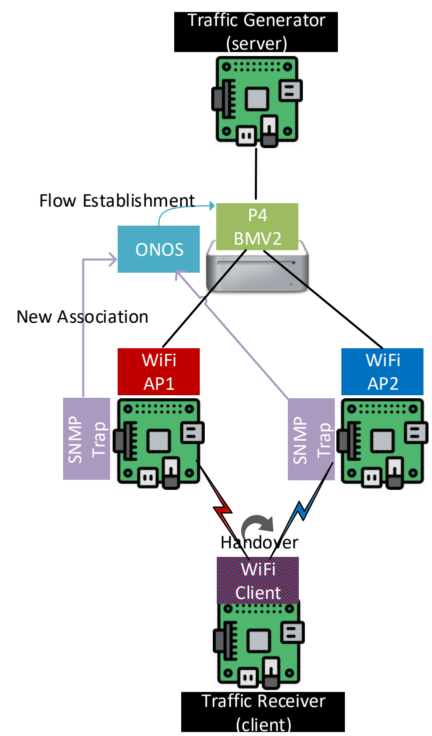

Fig. 3. Deployment for evaluation of the P4-based handover scheme

IV. Mobility Solution Evaluation

In this section we present a first evaluation of the mobility management scheme, using a scaled down version of the train use case. We use four Raspberry Pis (Rpi) for forming our network as follows: two of them are used as the home and target WiFi APs, one of them is used as the server in

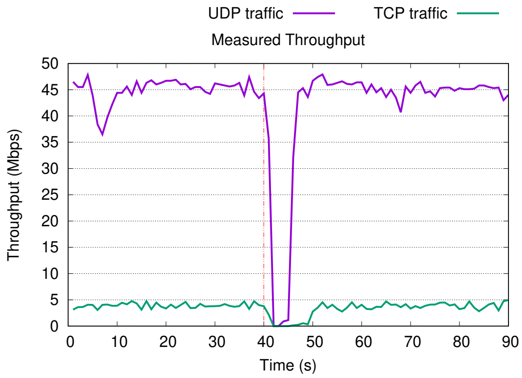

Fig. 4. Session Mobility Evaluation; red line denotes when the handover takes place. The session is preserved during the handover process without any drops in the end-to-end established session

the core network, and one of them as the user that roams between the two APs. Three of them are interconnected using a high-performance mini PC, with the P4 bmv2 switch running on top. ONOS is the controller managing the switch, and establishes the flows to/from the core network whenever a handover is happening. SNMP traps are running on the WiFi APs and send signals to the controller whenever a UE attaches on them. The overall topology is shown in Figure 3.

In order to test the network, we saturate the links with UDP/TCP traffic using the *iperf* traffic generator tool, the server side being located at the core network side and the client side on the UE. We focus mainly on preserving the established TCP connections throughout the process in order to test the feasibility of the scheme using the software switches. Figure 4 shows the collected experimental results. The experiment duration is 90 seconds, and at the 40^{th} second we instruct the UE to switch its AP association through external commands. As we see, for all the traffic cases we observe that there is a significant drop in the transferred throughput post the handover process. This is happening as the software switch struggles to establish the new flows. Although the messages from the SNMP trap interface are sent within one millisecond, and the controller assigns the flows instantly, the software switch is slow in establishing them as packet processing takes place entirely in application space. This is also reflected in the low TCP throughput that is less than 5Mbps; since processing takes place in software, additional delays are placed for packet processing, causing TCP re-transmission timers (RTO) to expire and therefore the congestion window to decrease. This fact is also reflected in a simple RTT analysis with the tool, as shown in Table I. Average RTT times for packets traversing the bmv2 switch are near 7 times higher than the ones compared with a simple kernel-based switch, like the one provided by the *bridge-utils* package. Nevertheless, since the focus of this work is to mainly test the feasibility of the session-continuity, the established sessions are not dropped during the entire experiment, allowing for a seamless handover. Moreover, the fact that the mobility pattern - along the tracks- is known, allows further lowering the handover times compared to common mobility patterns at cellular networks to just the time needed for the SNMP agents to transmit the information about the new association to the controller.

TABLE I

RTT Measurements

Solution	Avg.	Min.
bmv2	9.88 ms	8.16 ms
bridge-utils	1.44 ms	1.12 ms

V. Conclusion and Future Plans

In this work, we presented a mobility management framework that is currently under development for supporting a 5G railway use case field trial conducted by the 5G-VICTORI project. As the train crosses heterogeneous track-side deployments, management of the established sessions is needed for presenting the hosted on-board applications with a seamless end-to-end link. To this aim, a P4-compatible solution was designed, and evaluated using a scaled down version in a testbed environment. The results denote that session continuity is accomplished across the handovers, however end-to-end throughput results and latency are poor compared to other 5G-capable solutions. Since this result has been pinpointed to be caused by the software based implementation of the switch [10] due to packet processing taking place only in the application space, we foresee moving the solution to a hardware based P4 switch/smartNICs that allow higher switching speeds of several Gbps. The final handover scheme will be evaluated in-depth through network specific measurements, as well as application specific metrics for the deployed on-board services, measuring in detail the interruption time during the handover process.

References

[1] International Union of Railways, "Future Railway Mobile Communication System User Requirements Specification v5.0, 2020 [Online] https://uic.org/IMG/pdf/frmcs_user_requirements_specification-fu_7100-v5.0.0.pdf, Accessed Mar. 2021."

[2] "5G-VICTORI: Vertical Demos Over Common Large Scale Field Trials For Rail, Energy and Media Industries, [Online] https://www.5g-victori-project.eu/, Accessed Mar. 2021."

[3] "5G-PICTURE: 5G Programmable Infrastructure Converging Disaggregated Network and Compute Resources, [Online] https://www.5g-picture-project.eu/, Accessed Mar. 2021."

[4] C. Tranoris and S. Denazis, "Patras 5G: An open Source based End-to-End facility for 5G trials," *ERCIM NEWS*, p. 10, 2019.

[5] J. Zou *et al.*, "Europe's First 5G-Ready Railway Trial Utilizing Integrated Optical Passive WDM Access and Broadband Millimeter-Wave to Deliver Multi-Gbit/s Seamless Connectivity," in *2020 European Conference on Optical Communications (ECOC)*, 2020, pp. 1–3.

[6] N. Nikaein *et al.*, "OpenAirInterface: A flexible platform for 5G research," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 5, pp. 33–38, 2014.

[7] P. Bosshart *et al.*, "P4: Programming protocol-independent packet processors," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 87–95, 2014.

[8] "P4 behavioral model (bmv2) software switch, [Online] https://github.com/p4lang/behavioral-model, Accessed Mar. 2021."

[9] S.-Y. o. Wang, "High-speed data-plane packet aggregation and disaggregation by P4 switches," *Journal of Network and Computer Applications*, vol. 142, pp. 98–110, 2019.

[10] "P4 behavioral model (bmv2) software switch - Low throughput issues, [Online] https://github.com/p4lang/behavioral-model/issues/920, Accessed Mar. 2021."