5G-VICTORI: Optimizing Media Streaming in Mobile Environments Using mmWave, NBMP and 5G Edge Computing

Louay Bassbouss, Mehdi Ben Fadhel, Stefan Pham, Anita Chen, Stephan Steglich, Eric Troudt, Marc Emmelmann, Jesús Gutiérrez, Nebojsa Maletic, Eckhard Grass, et al.

HAL Id: hal-03789035

https://inria.hal.science/hal-03789035

Submitted on 27 Sep 2022

5G-VICTORI: Optimizing Media Streaming in Mobile Environments using mmWave, NBMP and 5G Edge Computing

Louay Bassbouss[1], Mehdi Ben Fadhel[1], Stefan Pham[1], Anita Chen[1], Stephan Steglich[1], Eric Troudt[1], Marc Emmelmann[1]
Jesús Gutiérrez[2], Nebojsa Maletic[2], Eckhard Grass[2,5],
Stefan Schinkel[3],
Annette Wilson[4], Sven Glaser[4]
Christian Schlehuber[6]

[1] Fraunhofer FOKUS, Kaiserin-Augusta-Allee 31, 10589 Berlin, Germany
[2] IHP - Leibniz-Institut für innovative Mikroelektronik, Im Technologiepark 25, 15236 Frankfurt (Oder), Germany
[3] PaxLife Innovations GmbH, Große Weinmeisterstraße 2, 14469 Potsdam, Germany
[4] Rundfunk Berlin-Brandenburg, Marlene-Dietrich-Allee 20, 14482 Potsdam, Germany
[5] Humboldt-Universität zu Berlin, Institut für Informatik, Berlin, Germany
[6] Deutsche Bahn AG, An der Welle 3, 60322 Frankfurt a. Main, Germany

Abstract. This paper describes an approach for optimizing media streaming services in mobile environments by utilizing the fifth-generation mobile network technology (5G), including millimeter Wave (*mmWave)* high speed data links, 5G Edge Computing and relevant state-of-the-art media streaming standards, such as *Network Based Media Processing (NBMP)* and *Server and Network Assisted DASH (MPEG SAND)*. Although the solution described in this paper can be applied to media streaming in various mobility scenarios, the focus of this work is on rail environments and allowing video-on-demand (VoD) catalogues (Mediatheks) to be made available to train passengers via regular VoD apps – even when trains are not connected to the Internet. To achieve this objective, we introduce caching nodes inside trains and their respective stations. The work presented in this paper will be trialed in a railway station in Berlin, as part of the EU-funded research project, 5G-VICTORI.

Keywords: 5G; mmWave; NBMP; Edge Computing; Data Shower; SAND; CDN

1 Introduction

The aim of digital technologies in the framework of the Fourth Industrial Revolution is to influence a large variety of sectors, with an emphasis on vertical markets for ICT industries, such as transportation, energy, media, and factories of the future. A

few examples stemming from this influence include the offering of connected goods, collaborative and automated processes within and across sectors, optimized processes (energy, transportation, logistics, etc.), and new and improved services concerning safety and security.

5G platforms play an essential role in bringing technology players, vendors, operators and vertical markets together. Their interactions are orchestrated in targeting new business models and opportunities for the ICT and vertical industries, and enabling cross-vertical collaborations and synergies in order to offer additional enhancements in value propositions.

There is a clear need to deploy 5G solutions for vertical industries. One of the first steps in the deployment process is the development of 5G infrastructures with flexible architectures in order to address a wide range of vertical applications. These infrastructures would offer converged services across heterogeneous technology domains and have a unified software deployment. In this context, the EU-funded research project, 5G-VICTORI [1], aims to conduct large-scale trials for advanced use case verification in commercially relevant 5G environments for transportation, energy, media and factories of the future, including specific use cases involving cross-vertical interaction. This objective is based on the work being developed within the 5G-PPP ICT-17 5G experimental platforms [2], 5G-VINNI [3], 5GENESIS [4], 5G-EVE [5] and 5GUK [6]. The Berlin cluster exploits the Berlin Platform, which is a 5G infrastructure that is currently being developed in the context of the ICT-17 5GENESIS Project. This cluster will run three use cases at the Berlin Central Station (*Berlin Hauptbahnhof*): 1) digital mobility 2) rail critical services and 3) CDN and media services in dense and mobile environments. This paper will focus on the work being developed within the third use case in order to optimize media streaming services in mobile environments enabled by 5G technologies, including mmWave [8], Edge Computing and media streaming standards such as NBMP [9] and MPEG SAND [10]. The details of this approach are described in the following sections.

2 Background and Concepts

This section presents the relevant components that play significant roles in optimizing media streaming with 5G: content delivery, monitoring and analytics, and media processing.

2.1 Data Shower and Multi-CDN

Streaming services usually deliver their content via content delivery networks (CDNs) **Error! Reference source not found.**, which are caches strategically placed at certain nodes of the Internet. These caches hold copies of content in order to

optimize media delivery for end-users. CDNs enable a higher level of service reliability and can be designed to optimize the distribution of network load and transmission costs. This concept is extended with 'Multi-CDNs [15]', which connects multiple CDN providers into one large network and combines the power of each. The fastest and most reliable CDN in a region is then selected. In 5G-VICTORI, we follow the multi-CDN concept and extend the streaming service's CDN to trains by equipping them with caches. These caches are filled with content via mmWave high-speed links or so called "Data Shower" to ensure seamless connectivity. This approach allows the transfer of media content from the CDN cache at the train station into the train's cache with multi-Gbps data rates. Data showers are installed at selected locations along the train route – with a starting point at train stations. The Berlin Central Station, a facility of Deutsche Bahn, is used to validate this use case for real-life environments by using the video-on-demand (VoD) catalogue, Mediathek, offered by a German public broadcaster, Radio Berlin Brandenburg (RBB).

2.2 Streaming Monitoring and Analytics via MPEG-SAND

To ensure the reliability of the end-to-end streaming workflow (encoding, packaging, delivery, playback, etc.), it is important to make the right instruments available for monitoring the entire workflow. For this purpose, we use the MPEG-SAND standard, specified in ISO/IEC 23009-5 [11], which defines standardized message formats for communication between server, client and network elements involved in the streaming process of MPEG-DASH. SAND-capable components can exchange real-time information of network and servers, as well as player behaviour and performance data. The SAND Metric Reporting will be integrated into the core of the 5G Network and ARD player (used in the RBB Mediathek), supporting shared resource allocation of multiple streaming clients that compete for bandwidth within the same network.

2.3 Network-Based Media Processing – NBMP

Leveraging processing capabilities and resources in a network is a trending approach in order to accomplish complex media processing tasks, while efficiently utilizing available resources and ensuring the potential for scalability. Deploying such complex services on top of microservices that may run on different cloud services, on an edge, or on-premise, can be a very intricate and time-consuming task.
Network-Based Media Processing (NBMP) [9] provides the high flexibility that is necessary for deploying such complex workflows on top of microservices. NBMP defines interfaces, media, and metadata formats to address fragmentation issues and

offer a unified way to perform media processing tasks on top of any cloud platform or IP network. For this reason, we considered the NBMP standard for the deployment of Data Shower and multi-CDN applications.

NBMP consists of three main network entities (see Fig. 1):

1. **Workflow-Manager:** the central entity that manages and configures NBMP workflows and relevant functions.
2. **Tasks**: implements one media processing function at a time. A *Function*, once loaded, becomes a *Task*, which is then configured by the Workflow Manager through the Task API and can begin to process incoming media.
3. **Function Repository**: offers APIs for function discovery and loading function description documents.
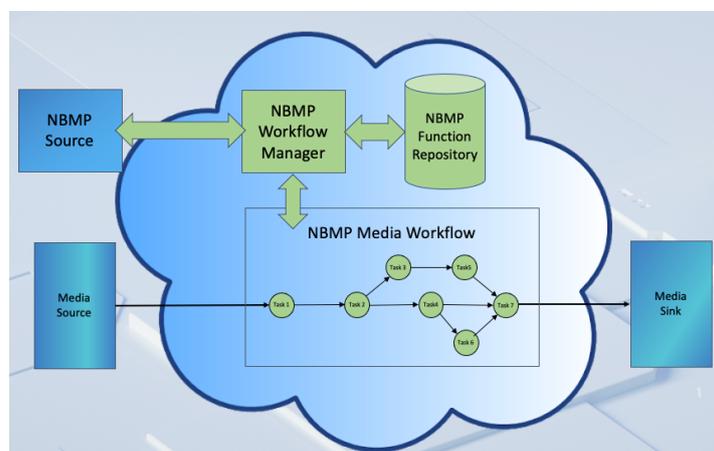

Fig. 1 NBMP Network Entities Overview [12]

NBMP allows the ability for easier deployments and service updates on distributed architectures, which opens the horizon for several use cases such as "network-assisted media quality enhancement" and "network-assisted media distribution".

3   Data Shower Architecture and Deployment

As mentioned in Section 2.1, the main approach followed in the 5G-VICTORI project (to optimize media distribution in mobile environments) is to extend the streaming service's CDN to train stations and trains by equipping them with caches. Media content is transferred from the station cache to the train cache via high-speed mmWave data links, during the time in which trains make their scheduled stops at each station. This installation is called a "Data Shower", which means that a large amount of data is transferred between two endpoints within a short amount of time. While connected to the train's onboard Wi-Fi, passengers can stream the content from

the train cache to their personal mobile devices without any interruptions and/or worries of high mobile data plan usage. A high-level overview of this approach is depicted in Fig. 2. In this diagram, there are two users of the same ARD Player App: ARD Player (1) retrieves media content from the CDN cache via public Internet, while ARD Player (2) retrieves content from the train cache while connected to the onboard Wi-Fi in the train.
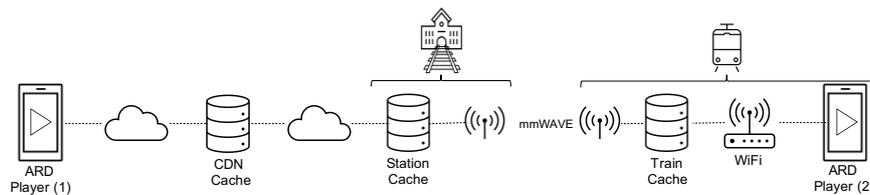

Fig. 2 High Level Overview

This approach has several advantages. First, content licenses that are already purchased by streaming services for conventional distribution via the Internet, are considered valid. It would be different if the streaming service offered its content via a portal of the train company. In this setup, the train company would become a third-party provider and thus, an additional entity in the exploitation chain. In our setup, the train company is merely the provider or host of the distribution infrastructure. The legal situation is therefore the same as with classic CDNs. A second advantage is that passengers can use the regular app of a streaming service (web or mobile) that is already installed on their personal device(s). Once connected to the train's onboard Wi-Fi, users will experience high-quality on-demand video without any strain on their mobile data budgets.

For cache nodes placed in trains, we rely on standard solutions that are mature in terms of performance and have proven themselves to be efficient in large-scale deployments, such as NGINX [14] . [14] To use NGINX as a for caching purposes, it is operated in a reverse proxy mode. Requests for which NGINX stores responses are then answered directly.

In most cases, not all content of a VoD service's catalogue will fit into the train's cache. High demands are placed on hardware installed in trains, for example, with regards to energy consumption, temperature development and reliability. Accordingly, the hardware is expensive. Content catalogues of large VoD services like ARD-Mediathek [13] have storage requirements of well over 100TB. Storage spaces of this size are not bargainable, even in the consumer segment. At the latest time in which several VoD services want to offer their services in trains, storage space becomes scarcer and each service must be selective about the content that is made available. Our system currently allows for an editorial preselection of content based on certain criteria. It is conceivable that this provision will later be

supplemented or replaced by an automatic selection system, which can foresee what passengers are keen to watch, such as recommendation engines.

The NBMP-enabled Data Shower architecture is depicted in Fig. 3. It describes the functional building blocks of the system that can be deployed on computing and storage nodes at train stations, in trains or in the cloud. Train station components are assumed to have a continuous Internet connection and are hosted on computing resources that are physically located close to the mmWave radio link equipment. This principal applies to components in trains as well, however, do not require a permanent Internet connection.
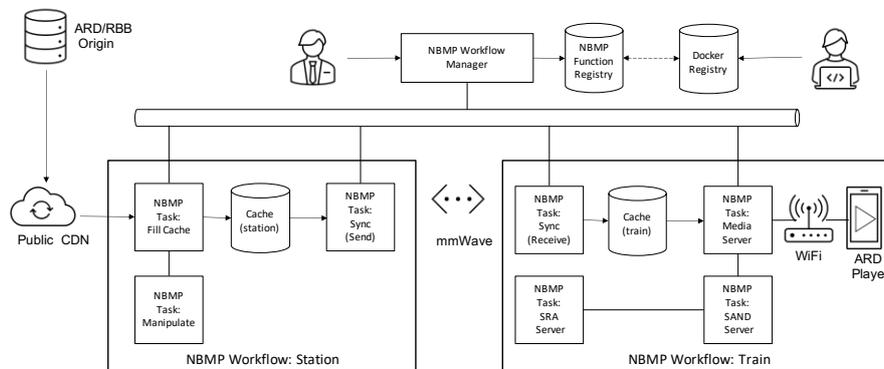

Fig. 3 NBMP Enabled Data Shower Architecture

The station and train components are deployed as NBMP workflows. Each workflow component is an NBMP Task, which implements a specific function that can be easily integrated into the workflow and communicates with other tasks through unified APIs and messaging mechanisms. The workflows can easily be defined, configured and updated via the NBMP workflow manager that is running in the cloud. In a proof-of-concept implementation, each NBMP function is implemented as a Docker image that is available in a centralized Docker registry. The NBMP function registry provides all necessary information in order to deploy a function that becomes an NBMP task. In our case, the task is a Docker container. As such, newly developed features can easily be accessible via the registry and used in new or existing workflows. This functionality becomes more vital in scenarios where media functions are provided by different vendors.

The trial will support the testing of communication between the mmWave unit installed at the rolling stock (Deutsche Bahn train) and the mmWave unit attached to the infrastructure (railway station). In such mobile environments, the use of directional antennas (together with beamforming / beam steering capabilities) becomes necessary to ensure seamless connectivity.

Fig. 4 sketches the scenario of the test trial at Berlin Central Station. As the train is approaching or departing the station, beam alignment mechanisms are required to avoid significant signal drops. However, frequent invocations of these mechanisms could lead to a deterioration in performance if the link is not continuously maintained.

The COTS mmWave devices used in the trial will operate at 60-GHz band (i.e., 57-66 GHz) in the four non-overlapping channels used in 802.11ad, where the center frequencies are 58.32, 60.48, 62.64, and 64.8 GHz. The channel bandwidth is 2.16 GHz and the maximal transmit power of approximately 23 dBm, allowing the device to reach a distance of 150 m. The devices allow for an achievable data rate of 2.5 Gbps.
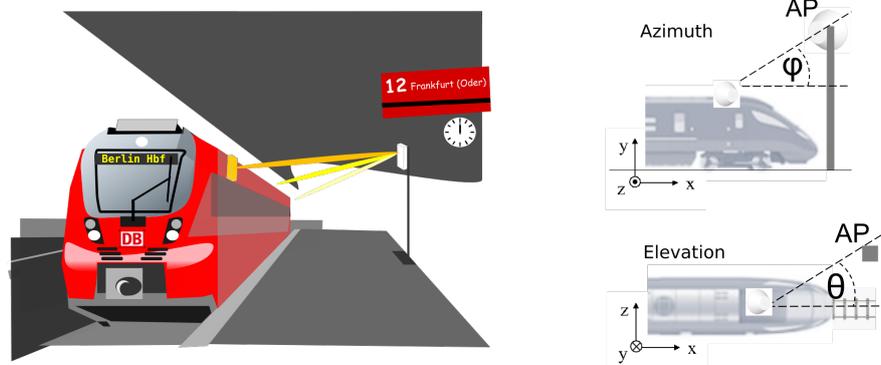

Fig. 4 a) Sketch of the track-to-train communication using mmWave pencil beams b) Azimuth and elevation requirements depending on the angles, COTS hardware used for the trial

The server installed on the train features two SFP+ connectors at 10 Gbps, which is connected to the mmWave unit attached to the upper part of the train. At the infrastructure level (platform), a simple ruggedized NUC with an IP68 rating is connected to the mmWave unit via a 10 Gbit Ethernet/SFP+.

For ensuring seamless connectivity between the two mmWave units, an analysis of the azimuth (XY plane) and elevation (XZ plane) angles between the mmWave unit attached to the train and that of the infrastructure must be carried out. The goal of this analysis is to determine the optimal mounting spots and orientation of both mmWave units. The speed of the train and height at which the units are installed, along with the rate of change of the azimuth and elevation angles as the train moves along the track, will be key parameters that will allow for an evaluation of the beam/sector changes and adaptation time offered by the equipment. This analysis will be required to ensure a maximal data transfer whilst both units are in reach, therefore allowing the transfer of big chunks of media data from the infrastructure to the train.

4   Conclusion

Streaming videos to train passengers can be a frustrating experience, especially if the Internet connection along the train route is not reliable. To address this issue, the 5G-VICTORI project follows the approach of extending conventional CDNs into caches that are deployed in trains and their respective stations, thus improving the overall user experience. We propose that, unlike conventional CDNs, cache nodes in the train do not need to have a permanent connection to the Internet but can be filled with media content via a wireless (mmWave) high speed data connection equipment installed in train stations and in trains. This method occurs shortly before and after the train stops at each station. In this paper, we described the approach, architecture and how state-of-the-art media streaming technologies and standards (such as NBMP and MPEG-SAND) are utilized to achieve our objectives. A planned field trial at Berlin Central Station using the RBB video catalogue will assess the benefits of providing video streaming for train passengers as they reach the stations, and if this concept can be adopted in other locations and means of transportation.

References

[1]   5G-VICTORI H2020 5G-PPP Project, https://www.5g-victori-project.eu/, last accessed INSERT DATE

[2]   5G PPP "The 5G Infrastructure Public Private Partnership" Infrastructure Projects, https://5g-ppp.eu/5g-ppp-phase-3-1-projects/, last accessed INSERT DATE

[3]   5G-VINNI "5G Verticals Innovation Infrastructure" ICT-19 Project, https://www.5g-vinni.eu/, last accessed INSERT DATE

[4]   5GENESIS "5th Generation End-to-end Network, Experimentation, System Integration, and Showcasing" ICT-17 Project, https://5genesis.eu/, last accessed INSERT DATE

[5]   5G-EVE "5G European Validation platform for Extensive trials" ICT-17 Project, https://www.5g-eve.eu/, last accessed INSERT DATE

[6]   5GUK Test Networks, University of Bristol 5GUK Test Network, http://www.bristol.ac.uk/engineering/research/smart/5guk/, last accessed INSERT DATE

[7]   Peng, G., "CDN: Content Distribution Network". CoRR cs.NI/0411069. (2004)

[8]   Niu, Y., Li, Y., Jin, D. et al. A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges. Wireless Netw 21, 2657–2676 (2015). https://doi.org/10.1007/s11276-015-0942-z, last accessed INSERT DATE

[9]   Network-Based Media Processing (NBMP) vision paper, https://mpeg.chiariglione.org/standards/mpeg-i/network-based-media-processing/nbmp-vision-paper, last accessed INSERT DATE

[10]   DASH-IF Position Paper: Server and Network Assisted DASH (SAND),: https://dashif.org/docs/SAND-Whitepaper-Dec13-final.pdf, last accessed INSERT DATE

[11]   ISO/IEC 23009-5:2017 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 5: Server and network assisted DASH (SAND), Online: https://www.iso.org/standard/69079.html, last accessed INSERT DATE

[12]   ISO/IEC 23090-8:2020 Information technology — Coded representation of immersive media — Part 8: Network based media processing, Online: https://www.iso.org/standard/77839.html, last accessed INSERT DATE

[13]   ARD Mediathek -RBB, Online: https://www.ardmediathek.de/rbb/, last accessed INSERT DATE

[14]   NGINX | High Performance Load Balancer, Web Server, & Reverse Proxy, Online: https://www.nginx.com/, last accessed INSERT DATE

[15]   V. K. Adhikari et al., "Unreeling netflix: Understanding and improving multi-CDN movie delivery," 2012 Proceedings IEEE INFOCOM, Orlando, FL, USA, 2012, pp. 1620-1628, doi: 10.1109/INFCOM.2012.6195531.