Arbitrating Network Services in 5G Networks for Automotive Vertical Industry

Original

Availability:
This version is available at: 11583/2801174 since: 2020-03-09T10:00:23Z

Publisher:
IEEE

Published
DOI:

Terms of use:
This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)
Abstract—This demonstration shows how the 5G-TRANSFORMER platform, and more specifically the vertical slicer, is capable of arbitrating vertical services. In this context, arbitration refers to handling the various services of a given vertical customer according to their SLA requirements, service priorities, and resource budget available to the vertical. In this demo, a low priority video service of the automotive vertical is terminated when a high-priority intersection collision avoidance service needs to be instantiated and there are not enough resources allowing all services to be run in parallel. All these services are deployed within the help of the 5G-TRANSFORMER platform in a multi-PoP scenario, with PoPs in Barcelona (Spain), Turin (Italy), and Pisa (Italy), featuring a high variety of transport and computing technologies.

I. INTRODUCTION

Network Function Virtualization (NFV) and network slicing have emerged as major 5G key enablers supporting the needs of multiple vertical services over a shared infrastructure. Thanks to such paradigms, a service can be deployed in the cellular infrastructure as a set of interconnected virtual network functions (VNFs), each demanding for a certain amount of computing, network, and memory resources, while guaranteeing the desired performance in terms of e.g., latency, reliability, and throughput.

To enable such vision, an efficient architecture has been proposed within the European 5G-TRANSFORMER (5GT) project [1], which consists of three main blocks. The Vertical Slicer (5GT-VS) is the entry point for verticals and provides a frontend through which a vertical can request a service and specify its composition in terms of VNFs and performance requirements, by filling in the Vertical Service Descriptor (VSD). The VSD is then translated by the 5GT-VS into a Network Service Descriptor (NSD), used for building the network slice supporting the requested network service (NS). The Service Orchestrator (5GT-SO) coordinates the end-to-end orchestration and manages the lifecycle of NSs, into which network slices are mapped. Finally, the Mobile Transport Platform (5GT-MTP) is responsible for managing the computing, storage and networking resources used to implement the VNFs, and the required transport resources to interconnect the different VNFs of a NS placed in different Points of Presence (PoP), as demanded by the 5GT-SO.

The resulting resource allocation should satisfy the service level agreements (SLA) expressed by the vertical in the service descriptors, while respecting the service priority and resource budget in terms of computing, storage, and network resources that the service provider (SP) must make available to match the fee paid by the vertical. We refer to this concept as arbitration [2]. Indeed, one of the most critical tasks for a SP is to deploy and maintain the vertical’s services, while meeting the corresponding SLA in spite of the highly variable service dynamics and associated traffic demand.

In this demonstration, we focus on the arbitration capabilities of the 5GT platform, which is managed by the arbitrator module of the 5GT-VS block. This block allows the 5GT platform to handle the different service requests by the same vertical while satisfying the agreed SLAs between the SP and the vertical. In brief, the objective of this demo is the validation of the 5GT architecture in facing these challenges, by considering an automotive vertical and two among the most relevant services for vehicular users, namely, vehicle collision avoidance at intersections (hereinafter referred as Extended Virtual Sensing (EVS) Service) and video streaming.

II. SYSTEM ARCHITECTURE

Fig.1 presents the setup used for the demonstration, which is split in three different geographical sites, namely Barcelona (Spain), Turin (Italy) and Pisa (Italy). These sites are interconnected by means of control and data plane virtual private networks (VPN). The whole 5GT platform (5GT-VS, 5GT-SO, 5GT-MTP and 5GT-MON) is placed in Barcelona.

The vertical interacts with the 5GT platform through the 5GT-VS User graphical user interface (GUI). The 5GT-VS receives the vertical requests and interacts with the 5GT-SO, which coordinates the end-to-end orchestration of the network services (NSs) with the 5GT-MTP based on the NS constraints (e.g., latency constraints) and the available advertised resources (compute, storage, network). In this setup, the 5GT-SO uses Open Source MANO (OSM) to handle the instantiation of VNFs at the underlying NFVI-PoPs. At any point during the NS lifecycle, the 5GT SP has knowledge of the different deployed NSs thanks to the available GUIs at each of the building blocks of the 5GT platform.

At the infrastructure level, the 5GT-MTP coordinates three different NFVI-PoPs, each deployed at one of the aforementioned sites. Each of these NFVI-PoPs is managed by its own Virtual Infrastructure Manager (VIM) based on an OpenStack
all-in-one Devstack deployment. Besides the VPN connections, these NFVI-PoPs are interconnected by means of a wireless (IEEE 802.11ac) ring of four elements, as depicted in Fig.1. The Wide Area Infrastructure Manager (WIM) is based on the Ryu open source SDN Controller. Finally, the setup is completed with an IEEE 802.11p Roadside Unit (RSU) physical network function (PNF) including a Radio Network Information Service (RNIS) to provide radio access to users (i.e., vehicles) and channel state information to applications, respectively.

III. DEMONSTRATION

We demonstrate the automated fulfilment of the established business SLAs between a vertical, namely an Automotive vertical, and the 5GT SP, by using the 5GT platform and the infrastructure in Fig. 1. A demonstration video is available online. The 5GT platform satisfies the SLAs in two ways. First, the 5GT platform deploys automatically the requested automotive NSs while meeting the constraints (e.g., latency) expressed in the corresponding network service descriptors (NSDs). Second, and this is the main focus of this demonstration, the 5GT platform satisfies the resource budget available to the vertical through the arbitration capabilities.

Two types of NSs are deployed during the demonstration, namely the Video Streaming (VS) NS and the EVS NS. The VS NS provides multimedia content to vehicular users, while the EVS NS enhances road safety by detecting and alerting vehicles approaching an intersections that are on collision course. The VS NS presents two deployment flavours, with different priorities, as specified at the moment of onboarding through the 5GT-VS. The "high priority" flavour consists of one video server VNF holding a web-server and a cache of videos, indicated as Video_Server_X in Fig. 1. The "low priority" flavour adds, to the above VNF, another VNF (Video_Control_X), which interacts with the RNIS to adjust the video streaming quality based on channel quality experienced at the vehicle Onboard Unit (OBU). This flavour of the VS NS has lower priority because it consumes more resources and the primary goal of the Automotive Vertical is to offer the video streaming service, no matter which video quality can be provided. The EVS NS consists instead of three VNFs, namely the Cooperative Information Manager (CIM) DB, the Warning message VNF, and the EVS VNF. A full description on the logics of this NS can be found at [3]. Being a safety service, the EVS NS has high priority. The main steps of the demo are as follows:

1) The Vertical User requests the instantiation of the two VS NS flavours through the 5GT-VS User GUI. The 5GT-SO processes the request as explained in [4], deploying the VS NSs in the underlying infrastructure while satisfying the NS constraints embedded in the descriptors. The "low priority" instance of VS NS is deployed between two PoPs and its "Video_Control_2" VNF is placed close to the RNIS as it needs to leverage low-latency channel quality measurements.

2) The Vertical User requests the instantiation of the EVS NS through the 5GT-VS User GUI. The arbitrator module of the 5GT-VS checks the resource budget available for this Vertical user and realises that the three NSs cannot run simultaneously. Based on the priority of the current deployed NSs and the new incoming request, the arbitrator decides to terminate the "low priority" VS NS to release resources for the EVS NS.

3) Once the 5GT-SO confirms the termination of the "low priority" VS NS, the 5GT-VS proceeds with the instantiation of the EVS NS, which is deployed in the PoP close to the RSU to honor the latency requirements embedded in the request, as depicted in Fig. 1.

REFERENCES


Available at https://youtu.be/QKQyZ0I0Nto