Li, X., Deiss, T., Mangues-Bafalluy, J., Baranda, J., Costa-Perez, X., Landi, G., Bernardos, C. J., Iovanna, P., Zurita, A. & Bertin, P. (2020). Automating Vertical Services Deployments over the 5GT Platform. *IEEE Communications Magazine, 58*(7), 44–50.

DOI: 10.1109/mcom.001.1900582

Automating Vertical Services Deployments over the 5GT Platform

Xi Li*, Thomas Deiß†, Josep Mangues-Bafalluy‡, Jorge Baranda ‡, Xavier Costa-Pérez*, Giada Landi§, Carlos J. Bernardos‖, Paola Iovanna¶, Arturo Zurita** and Philippe Bertin††

*NEC Labs Europe, Germany, †Nokia Solutions and Networks, Germany, ‡CTTC, Spain, §Nextworks, Italy, ¶Ericsson Research, Italy, ‖ University Carlos III of Madrid, Spain, **ATOS, Spain, ††Orange, France

Abstract— **This article presents a system for 5G networks that allows to meet diverse needs of vertical industries simultaneously sharing the same physical infrastructure. *Orchestration*, *Network Slicing*, *Edge Computing*, and *Federation* are key technologies enabling industry verticals to have their own *virtual networks*, which might require aggregating transport networking and computing fabric, from the edge up to the core and cloud. Three novel building blocks are defined to meet these challenges in an automated manner: (*i*) a *Vertical Slicer* as the entry point to create services and request slices, (*ii*) a *Service Orchestrator* to manage the services and decide their placement and allocation of required resources and (*iii*) a *Mobile Transport and Computing Platform* virtualizing infrastructure networking and computing resources in an integrated manner. An experimental evaluation of the developed system shows its feasibility and confirms some of the benefits expected.**

I. Introduction and motivation

Network slicing is the main concept enabling 5G networks to serve multiple vertical industries (e.g., Industry 4.0, automotive or emergency services) over a common infrastructure, meeting their individual diverse and stringent service requirements (like ultra low latency, high reliability and large capacity), while decreasing CAPEX and OPEX of the network by allowing infrastructure sharing among multiple tenants.

Towards this goal, the 5G-TRANSFORMER project (referred as 5GT) has worked on developing an evolved 5G mobile transport network platform based on Network Function Virtualization (NFV) and Multi-access (Mobile) Edge Computing (MEC), allowing to deploy vertical services in dedicated or shared network slices. Moreover, each slice may span across several data centers, RAN (3G/4G and 5G), different transport technology domains and multiple administrative domains.

5GT focuses on a slice-specific challenge: the need to expose the capabilities of the network, offer control to the verticals and allow automated service creation, provisioning and management. Our main contributions are:

- (*i*) presenting the overall design of the 5GT system architecture, describing in detail the key functionalities of individual building blocks along with their internal functional components and enabling technologies and implementation solutions used herein. This extends our work in [1], where the basic concept and guidelines for a high level system design were described;
- (*ii*) introducing main capabilities of the 5GT platform for supporting various vertical services. The 5GT solution has been published as open-source software [2];
- (*iii*) initial experimental evaluation of the open source 5GT solution comprising a real vertical service example and demonstrating the advantages of the 5GT platform to enable automatic service provisioning from service creation up to service instantiation and deployment.

The rest of the article includes the related work in Section II followed by a detailed description of the 5GT system design (Section III). Proof-of-concept results are provided in (Section IV). Finally, Section V presents our conclusions.

II. Related Work

Nowadays, in practice most of operators' networks provide almost none or only limited functionality to support verticals to allow them directly onboard and deploy their services over the operator's infrastructure. To this aim, different 5G projects [3] are working on different network slicing solutions based on SDN/NFV. However, they are mainly focused on the level of network services and infrastructures from operator point of view, but not providing means for verticals to automatically create and provision their services without knowledge of underlying networks. To lower the entry barrier for the verticals, 5GT proposes a catalogue of service definitions, called *blueprints*, which can be completed with actual values regarding dimensioning and SLA requirements during service onboarding phase. We expect that many vertical services can be predefined in such a catalogue, which are offered to verticals through a well-defined northbound interface to enable them to specify service requests.

The GSMA [4] defined the notions of generic slice template (GST), and the instances of them, network slice templates (NEST). A GST defines attributes such as supported throughput, functionality, or APIs that characterize any slice, whereas a NEST provides specific values for these attributes. Service descriptions in 5GT also contain SLA requirements, but are more general than GSTs and NESTs as they also take the vertical applications within vertical services into account. In [5], differences in network slicing views from GSMA and other SDOs are analyzed to propose a unified network slicing model, which in turn is mapped to the ETSI MANO architecture. Compared to this, 5GT implemented the proposed service blueprint model as described in Section III-A, leveraging on these SDO models.

For service orchestration, various orchestration platforms have been developed, e.g., OSM, Cloudify, ONAP, to automate the deployment and configuration of Network Services. Initial benchmarking of OSM and ONAP in [6] confirms their capability to successfully on-board and instantiate VNF and NS instances. [3] also summarizes different orchestrators and enabling technologies for 5G networks. Despite the growing adoption of these platforms, we identified many challenges to support a vertical service deployment and orchestration solution meeting vertical specific requirements. First, they currently offer only limited support on network slicing as preliminary efforts are still at an early stage. Second, automated orchestration of multi-data center interconnection through a WAN (Wide Area Network) is not yet implemented. Third, none of the current orchestration platforms support orchestrating MEC services. Finally, so far there are almost no implementations of orchestration platforms supporting service federation mechanisms, which allows deploying a segment(s) of a service in another administrative domain owned by another operator. The 5GEx project worked on multi-domain multi-technology orchestration mechanisms, considering both resource and service aspects. However, its reference implementation focused mainly on resource orchestration for validation purposes. Besides, due to the different timings of 5GEx and 5GT projects, the 5GT solution is aligned with current standard ETSI NFV IFA specifications with some extensions and has been implemented supporting main Open Source MANO frameworks like OSM and Cloudify.

Towards the infrastructure level, we rely on the already developed 5G-Crosshaul [7] solution virtualizing fronthaul and backhaul networks integrating heterogeneous transport technologies and various computing resources. 5GT extends the 5G-Crosshaul solution to enable integration of MEC and RAN, dynamic placement of VNFs in the data plane, providing multi-level and multi-criteria abstraction and network control over designed APIs.

III. 5G-TRANSFORMER ARCHITECTURE DESIGN

The 5GT system architecture extends the ETSI MANO design with new functional building blocks, namely the Vertical Slicer (5GT-VS) and the Service Orchestrator (5GT-SO) interworking with the Mobile Transport Platform (5GT-MTP), as shown in Fig. 1. Their key functionalities and the workflow of service instantiation upon receiving a vertical service request are also explained in Fig. 1 and described in the following.

5GT architectural components interact to provide horizontal and vertical elasticity to adapt to diverse service requirements, dynamic network changes and heterogeneous infrastructures. In this sense, 5GT embeds mechanisms for (*i*) service arbitration to handle inter-service and inter-slice interactions as well as slice sharing; (*ii*) service scaling to adapt to service changes and resource availability; (*iii*) multi-domain orchestration to provide inter-technology domain interaction and connectivity; and, (*iv*) service federation to dynamically request services to peering administrative domains and orchestrate E2E services and resources.

A. Vertical Slicer

The 5GT-VS is the common logical entry point (i.e., one stop shop) for all verticals to request the creation of respective slices to deploy their services. It offers vertical services through a high-level interface, allowing verticals defining services from a catalogue of vertical service blueprints (VSB), which, focus on specifying service logic with composed vertical applications and the requirements of vertical services. The vertical can select the required service from this catalogue and customize it by defining a vertical service descriptor (VSD) with additional service-level details (e.g., dimensioning of the service, required level of isolation, its coverage area, IP addresses for management systems, etc). In this way, all the technical details about how to deploy the service (e.g., type and size of VNFs, external connection points of the service, etc.) are completely hidden for the verticals. For instance, a RAN can be abstracted as a physical network function (PNF) and a service access point (SAP) with parameters describing the coverage area and radio access technology [8]. One could set the parameter layer Protocol of the SAP representing 5G air interface to ”New Radio (NR)”.

In the next step, the 5GT-VS translates service requirements into slice-related requirements, decides whether the service can be provided over an existing network slice or a new one needs to be established (depending on the requested level of service isolation), and manages the lifecycle of the network slices. In our system, the definition of network slices is aligned with the model from 3GPP [9], i.e., through a number of NFV network services (NFV-NS) that are instantiated and dynamically adjusted according to the slices’ composition and characteristics (derived automatically from the vertical service requirements). Vertical services with similar characteristics, no isolation requirements and belonging to the same tenant can share the same network slice. Moreover, concurrent network slices can share some components referred as network slice subnets.

More specifically, when the vertical user requests the instantiation of a new vertical service instance (VSI), this is mapped by the 5GT-VS into a network slice instance (NSI), which is realized through a NFV-NS instance (NFV-NSI) in the 5GT system and defined through a set of NFV-NSs using the NFV defined network service descriptors (NSDs) as defined in ETSI NFV IFA014. NSDs specify the information elements for a NFV-NS including the requirements for Virtual Links (VLs) and Virtual Network Functions (VNF), it also provides a set of Deployment Flavors (DFs) to allow variations in deployments of services with different parameter setting and number of VNF instances. In particular, the 5GT-VS first translates the VSD into a corresponding NSD that defines the structure and the size of the target slice, using translation rules to map service level requirements and cardinalities into deployment flavors and scaling rules to be used by the 5GT-SO. Eventually, the 5GT-VS decides how to compose the target slice (potentially re-using components from existing slices) and requests the 5GT-SO to create or update the associated NFV-NSIs. The core of the 5GT-VS includes three components providing its procedural and algorithmic logic.

The *VSI/NSI coordinator and lifecycle manager* handles the delivery and runtime of vertical services, coordinating their lifecycle based on verticals’ requests, e.g., to instantiate a new service, or scaling in response to monitoring notifications. This triggers corresponding actions at the respective

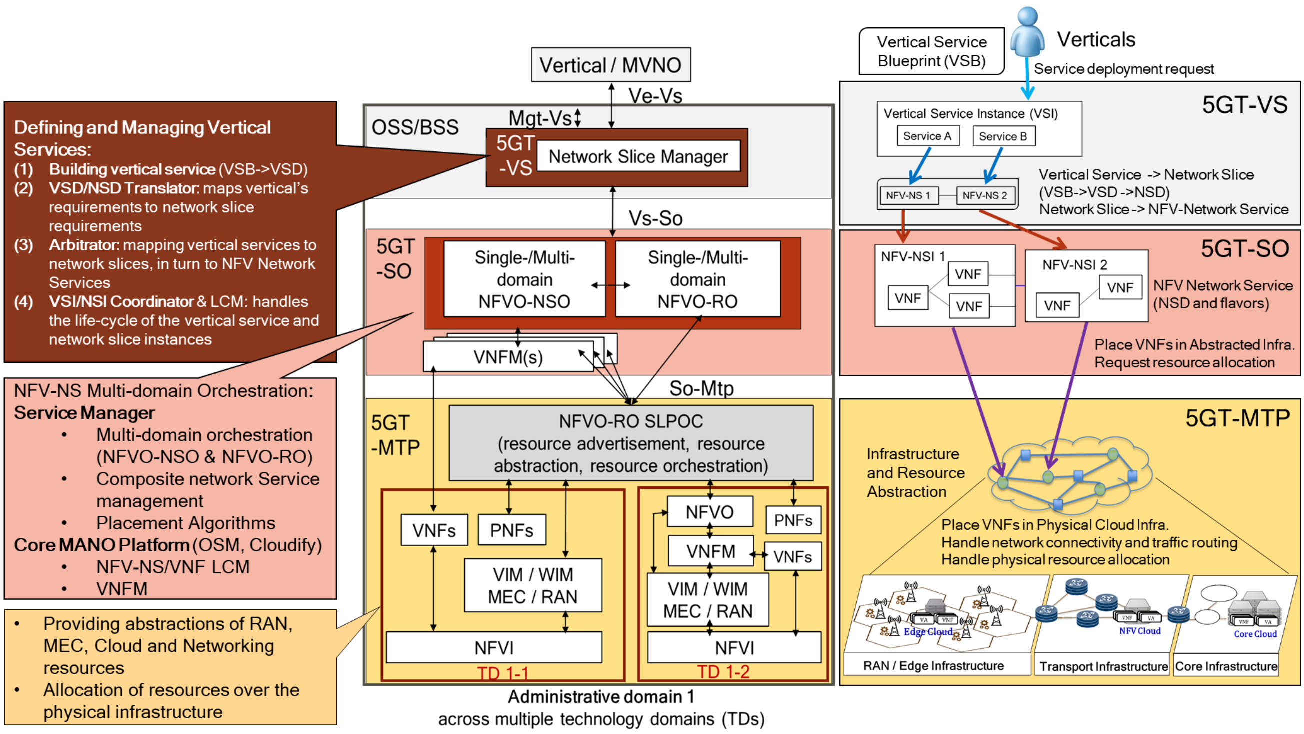


Fig. 1. 5G-TRANSFORMER System Architecture and Main Functionalities.

NSI and network slice subnet instances (NSSI), which in turn request the instantiation, scaling, modification or termination of the related NFV-NS instances at the underlying 5GT-SO.

The *VSD/NSD translator* handles the translation between VSDs and NSDs, mapping the verticals business oriented requirements into technical specifications focusing on network and resource oriented parameters needed by the 5GT-SO to perform a NFV-NS deployment. More specifically, based on the input from the VSB and values for the service parameters specified by the vertical in the VSD, the translator uses rules to select one or more NSDs and decide potential DFs and Instantiation Levels (IL) for the VNFs and the VLs referenced in the VSD, to ensure the network service able to meet the service requirements expressed in the VSD. Similarly, the translator can map SLA requirements to MANO policies for the 5GT-SO, describing how to optimize VNF placements considering latency or coverage constraints.

The *Arbitrator* controls access to a vertical's resource budget among vertical services belonging to the same vertical, and decides how the services are mapped to network slices, allowing dedicated slice or multiple services sharing common slices. The amount of needed resources for a VSI depends on placement decisions for VNFs and VLs, which are done by the 5GT-SO in later deployment phases. Therefore, the Arbitrator pre-calculates several DFs for different placement options with a range (e.g., minimum and maximum) of amount of resources [10] and then passes them to the 5GT-SO, which decides actual VNFs placement and which DF to use according to actual resource availability.

B. Service Orchestrator

The 5GT-SO is to provide an end-to-end orchestration of the NFV-NS requested by the 5GT-VS. These services may span across one or multiple administrative domains and they are deployed by interacting with the local 5GT-MTP and/or with other 5GT-SOs, respectively. Deploying different parts of the NFV-NS in different administrative domains is referred to as federation. In this sense, NFV-NSs may be (de)composed (e.g., a composite NFV-NS having multiple nested NFV-NSs) and request the deployment of each service segment to a peer 5GT-SO in charge of management and orchestration of another administrative domain. This is done in a transparent way to the 5GT-VS (and to the verticals).

The 5GT-SO receives networks service request in the form of an NFV-NSD, which describes a NFV-NS that the 5GT-SO should deploy (either on its own or by leveraging neighboring 5GT-SOs). The 5GT-SO has two main components, namely the service manager (SM) and the core MANO platform.

The *service manager (SM)* is the brain of the 5GT-SO, in charge of managing all NFV-NS databases (e.g., service catalog, resources, instances) and dispatching tasks to the rest of building blocks. It also includes the service orchestration engine (SOE), which corresponds to the role of NFVO-NSO as defined in ETSI NFV and it handles service orchestration tasks (including service federation), and the resource orchestration engine (ROE), which corresponds to the role of NFVO-RO as defined in ETSI NFV and it is aware of the resources abstracted by the underlying infrastructure layer (5GT-MTP) as well as the requested resources by the service. The resource orchestrator sends all this information to the placement algorithms (PA) [11] for deciding optimum function placement and resource needs. Based on the outcome of this computation, it instructs the 5GT-MTP where to place the functions in the virtual infrastructure and also allocates the required compute, storage, and networking resources through the southbound interface. In brief, the SM manages the lifecycle of the requested NFV-NSs, and it follows a modular design to integrate multiple *MANO platforms*.

The *core MANO platforms* (e.g., OSM, Cloudify) are integrated into the 5GT system through wrappers. Such design provides more flexibility for service deployment (e.g., each domain is free to use a different MANO platform) and cope with individual MANO platform evolution. Such MANO platforms orchestrate the cloud resources, which are stitched with the corresponding network resources at the SM, resulting in an operational E2E NFV-NS.

Additionally, as part of NFV-NS lifecycle management, the SM, based on the rules specified in the NSD, carries out automatic network service management (e.g., scaling) to adapt NFV-NS deployments to service demands and network changes and to react in case of SLA violations. For this purpose, the 5GT-SO also interacts with the monitoring platform (5GT-MON) by configuring monitoring jobs to measure relevant metrics at the VNF level and by handling alerts to trigger self-adapting actions. Moreover, the SM also supports updating network services requested by the verticals or the 5GT-VS. As a general rule, whenever there is an alter about the service, there is coordination between the SOE and the ROE logic at the 5GT-SO to instruct the 5GT-MTP, which eventually updates the resource allocation so that the E2E service is correctly deployed. Thereby, scaling elasticity is handled by both the 5GT-VS and the 5GT-SO, as they receive monitoring data and react to preserve SLAs. In the current implementation, the possible (DF, IL) pairs and auto-scaling rules are defined when the vertical service (and the associated NSDs) are designed according to the nature and requirements of the vertical service. This knowledge is integrated in the translation and arbitration operations in the 5GT-VS, which selects the NSDs with the appropriate scaling rules as a result of its vertical service-to-network service mapping. More specifically, at the 5GT-SO, the scaling workflow works as follows. The NSD of the deployed NFV-NS states the metrics to be monitored by the 5GT-MON (configured during instantiation). Furthermore, the scaling rules in the NSD define the conditions that these metrics must meet to trigger an alert and they are also configured in 5GT-MON. Upon an alert from 5GT-MON, the 5GT-SO will take the corresponding scaling action (e.g., if CPU consumption is above 75% for more than 20s, switch from one (DF, IL) pair to another one adding more VNF instances to balance the load of the service). Deploying a new VNF instance as a consequence of scaling out action results in: (*i*) the 5GT-SO sending a request to the 5GT-MTP to deploy a new virtual machine (VM) embedding the required VNF functionality including the VM image previously on-boarded; (*ii*) the 5GT-MTP to deploy the VM and attaching it to the appropriate intra-PoP according to the structure of the selected NSD; (*iii*) afterwards, the 5GT-MTP configuring the transport network connections that connect this new VM with the rest of VMs in the service (which may be already deployed in other NFVI-PoPs); and (*iv*) associating the monitoring jobs and alerts to this new VNF to enable auto-scaling operation of the VNF following the scaling rules.

As for federation, both service and resource management building blocks are integrated in the SM to interact with their peers in the SM of the 5GT-SO of the other administrative domain. At a high-level, the SOE is in charge of distributing the nested NSs among the local and peering domains by generating the corresponding ETSI NFV IFA013 requests, which carry some optional information elements (e.g., AdditionalParamsForNs). At the resource orchestration level, there is a block specifically designed to coordinate the assignment or resource IDs (e.g., IP addresses, VLAN IDs) by interacting with the same block of the peering domain and the underlying 5GT-MTP. This guarantees E2E resource orchestration in multiple domains.

C. Mobile Transport and Computing Platform

The 5GT-MTP manages the underlying physical mobile, transport and computing infrastructure, responsible for deploying VNFs over the infrastructure and providing resources required by the 5GT-SO to deploy the NFV-NSs. Essentially, it handles the VNF placement and instantiation on the physical cloud infrastructure and setting up the network connectivity, managing the allocation of actual physical networking and cloud resources to deploy the service. It also hides the complexity of the underlying infrastructure by exposing the 5GT-SO a desired abstraction.

The main components of the 5GT-MTP are: *Virtual Infrastructure Manager (VIM)* that manages storage, networking and computational resources in its respective NFVI-PoP domain, *WAN Infrastructure Manager (WIM)* that provides inter-domain connectivity, *MEC Controller* that configures and manages MEC-specific services at the MEC platforms/hosts, *RAN Controller* that controls the radio domain as well as their connectivity to the mobile core, and *Network Function Virtualization Infrastructure (NFVI)*.

The NFVO-RO is the single logical point of contact. It has the full view of the resources of each technology domain, and upon receiving a resource allocation request from the 5GT-SO it forwards the corresponding request to the relevant control entities (e.g., VIM or WIM), allocating virtual resources needed to deploy the VNFs that form the NFV-NS, interconnect them, and/or configuring relevant parameters of the PNFs. As for the interconnection, once the inter-PoP logical link is selected at the 5GT-SO, the request reaches the 5GT-MTP, which calculates the actual path based on the topology information exposed by the WIMs. The corresponding requests to set it up are sent to the WIMs, which eventually generate flow rule setup requests towards each SDN controller involved in the path.

The 5GT-MTP can decide which level and granularity of resource abstraction exposed to the 5GT-SO. In TD 1-1, the 5GT-MTP exposes to the 5GT-SO the available virtual resources through the different controllers and/or PNFs, like physical eNBs or pGWs to allow direct configuration. In TD 1-2, the 5GT-MTP provides an abstract view that is composed by NFVO. The demarcation points for a 5G RAN could be at the interface in the user plane between the core network (CN) and the RAN, named N3, or the interface between the CN and an external data network, named N6. In [8], we consider N6 as demarcation point. All the RAN equipment from the antenna to the User Plane Function (UPF) interface N6 is abstracted as POP, named Radio POP. This is a virtual POP including radio equipment and the transport connections in the RAN. Geographical information is associated to such POP to indicate the location where the RAN functions can be provided. Additional parameters can be provided such as bandwidth and latency. This solution allows to expose to the 5GT-SO a similar abstraction view for the RAN as for a traditional POP. This was demonstrated in the eIndustry proof of concept, see [12].

IV. Experimental Proof-of-Concept

This section validates the defined 5GT architecture to automatically deploy network slices serving vertical services. We present the characterization of service creation time (SCT), one of the 5G key performance indicators, for a specific vertical service, namely virtual Content Delivery Network (vCDN), defined by the 5GT vertical partner providing entertainment and media services.

A. Description of NS under evaluation and benchmarking

Fig. 2 depicts the vCDN service, which consists of three different VNFs: webserver, cache server and origin server. The user interacts with the webserver to demand the video content. Then, the cache server VNF returns this content directly to the webserver if available or asks the origin server the requested content. This vertical service presents two different instantiation levels (ILs), differentiated by the amount of deployed cache server VNFs (one versus two), so the 5GT platform would deploy one IL or the other depending on the requirements expressed by the Vertical User when filling the VSD (e.g., expected peak number of users to serve). Despite the fact that an increasing number

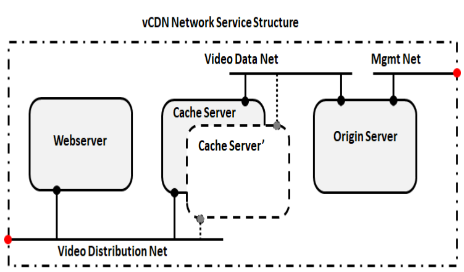


Fig. 2. vCDN Service under evaluation, showing its different ILs.

of companies are virtualizing some of their IT components and that there is research work (e.g., [13]) which automates some parts of the deployment process, we took manual operation as initial reference for comparison based on our discussions with operators and verticals. See [13] for a detailed discussion. In traditional CDN deployments, all CDN components are usually deployed in bare metal hosts. This requires the installation of a dedicated physical server for each component. Their number varies based on the dimension of the CDN that, in turn, depends on the number of expected media users. A careful dimensioning of the CDN system is thus initially required. Its outcome is a CDN design that specifies the number, size and positioning of the caches, the capacity of the interconnections between the origin server and the caches, etc. By following this process, dimensioning plus installation and configuration of the system takes around 13 hours per appliance. The vCDN allows reducing the usage of hardware resources, due to the sharing of the physical infrastructure. Moreover, the automated orchestration of its service deployment – enabled by the 5GT system – allows to drastically reduce the service provisioning time from hours to a few minutes, as presented in Section IV-C. Furthermore, we also compare the results obtained with previous work in equivalent (though less complex) scenarios in which dynamic deployments are also in place [13].

B. System Prototype and Evaluation Setup

Our experimental evaluation of the 5GT system [2] is performed over the developed 5GT testbed. This testbed contains a set of servers featuring an Intel Core i7 processor (12-Cores at 3.3 GHz x86 CPU) with 32GB of RAM memory and 1TB of hard disk running the Ubuntu 16.04 LTS distribution. These servers can be dynamically connected to form diverse network topologies interconnected by a variety of transport technologies (optical/wireless). One of these servers hosts the 5GT system prototype. It includes the 5GT-VS to create the vCDN service and related slice upon the request of the vertical customer. More specifically, the VSB allows the media provider to select the number of desired CDN users, and then the 5GT-VS allocates a network slice customized for these requirements in the form of NFV network service. It then interacts with the 5GT-SO to automatically trigger the network service instantiation procedures. The 5GT-SO features a complete service manager that integrates OSM as MANO platform. The 5GT-MTP manages other servers of the experimental infrastructure running OpenStack-based VIMs and an ABNO-based WIM, using the COP protocol (Control Orchestration Protocol), a precursor of the ONF transport API, respectively. The 5GT-SO, together with the 5GT-MTP carry out the deployment of the VNFs and establish the connections among them. More specifically, based on the output of PA, the 5GT-SO instructs the 5GT-MTP to create the necessary intraPoP networks at the different VIMs where reside the VNFs of a NFV-NS to support the virtual links expressed in the NSD. Then, after the VNFs are deployed (interaction MANO platform with 5GT-MTP), the 5GT-SO requests to the 5GT-MTP the allocation of resources in the logical links interconnecting the different NFVI-PoPs to provide E2E connectivity between VNFs. These transport network resources are allocated thanks to the interaction between 5GT-MTP and WIM entities.

TABLE I

DESCRIPTION OF MAIN SERVICE CREATION TIME COMPONENTS.

Component	Description	Labels
5GT-VS	Initial processing when receiving request from vertical and final polling to check successful deployment of NFV-NS	VS IL-2 or VS IL-1
5GT-SO RO	NSD parsing, resource retrieval from 5GT-MTP, interaction with PA (request+calculation)	RO IL-2 or RO IL-1
5GT-SO+MTP: Create IntraPop Networks	Interaction between 5GT-SO wrapper and 5GT-MTP (incl. Openstack) to create intra-PoP network to which VNFs attach	Net IL-2 or Net IL-1
5GT-SO+-MTP: Allocate VNFs	Interaction 5GT-SO - 5GT-MTP for deployment of VNF	VNF IL-2 or VNF IL-1

C. Evaluation of Service Creation Time

This section evaluates the main components of SCT for the deployment of the two instantiation levels defined for the vCDN service in a single administrative domain. We define SCT as the elapsed time between the Vertical User launches the service instantiation at the 5GT-VS until the 5GT-VS declares this service as instantiated. Fig. 3 presents a comparison of the average values of the total SCT (total column size) as well as the main components of SCT for both instantiation levels. These components are described in Table I.

Fig. 3 shows that total SCT is in the order of few minutes for both instantiation levels, hence decreasing the required time compared to the manual case, previously explained. More specifically, the average total times are 83.8703s and 106.7803s for IL-2 and IL-1, respectively. This time is heavily correlated with the number of deployed VNFs, being its booting up process, the most time consuming operation in the instantiation process (61.9591s and 81.0637s for IL-2 and IL-1, respectively). The second most time consuming operation is the creation of intra-PoP network to which VNFs attach. Basically, these two operations correspond to the actual resource allocation operations at the physical infrastructure. Hence, operations corresponding specifically to the 5GT platform have limited impact on the SCT.

Previous work in equivalent scenarios resulted in similar deployment times (i.e., 299.52s for a 4 VNF service) [13]. At the MANO level, the scenario is less complex in the sense that the 5GT MANO stack adds new functionality, hence new logic to be executed, that is not present in previous work. Such additional logic is distributed in the various building blocks and includes, for instance, resource abstraction, multi-PoP support, E2E placement, scaling, NFV service composition, multi-administrative domain service federation, inter-vertical and inter-slice arbitration, or vertical service-to-NFV NS service translation. It is worth mentioning that despite all this additional functionality, SCT is below (or equivalent to) that of previous work for service deployments of equivalent or more complexity.

D. Service Creation Time Profiling

Figs. 4(a) and 4(b) present the average, maximum, minimum, median, and 20th and 80th percentiles of each of the components for both instantiation levels. They are separated into two figures for the sake of readability due to the magnitude of VNF allocation times and total SCT (Fig. 4(b)) in comparison with the rest of operations (Fig. 4(a)).

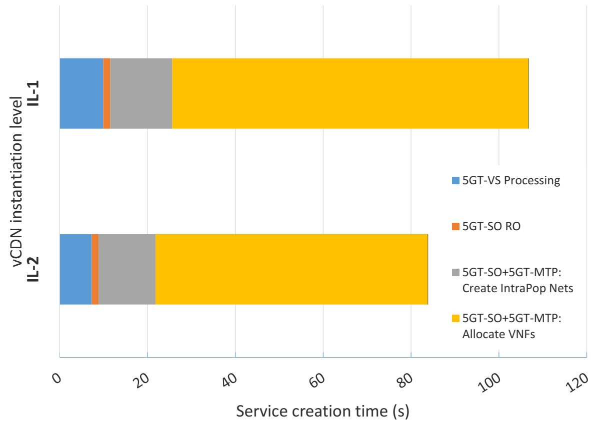


Fig. 3. Comparison of the main components of service creation time for two vCDN instantiation levels.

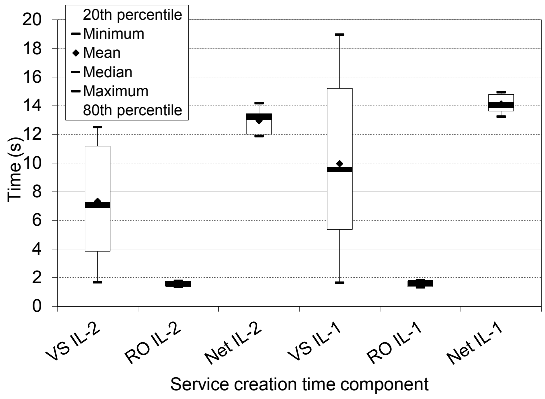
These figures show that the highest variability comes from the 5GT-VS processing component, although this operation is not the most time consuming with respect to the overall experienced SCT. This is due to the periodic polling operation (every 20s) that the 5GT-VS does to receive confirmation from the 5GT-SO that the NFV-NS has been successfully deployed. Since the rest of components show much lower (or at most similar) dispersion of values, this one also becomes the cause for most of the value dispersion for total SCT. For components involving internal processing inside the 5GT-SO, the values are quite similar, also because they are very small compared to the rest. For the biggest components, those corresponding to the IL-1 are always higher than those of the IL-2, though variability of the values does not necessarily increase.

The results presented in this section have shown the automation of the 5GT platform to deploy vertical services in the context of a specific vertical industry (media distribution). However, the 5GT system has been evaluated and demonstrated with several additional vertical industries: Automotive, E-Health, Smart Factories and MVNOs [12].

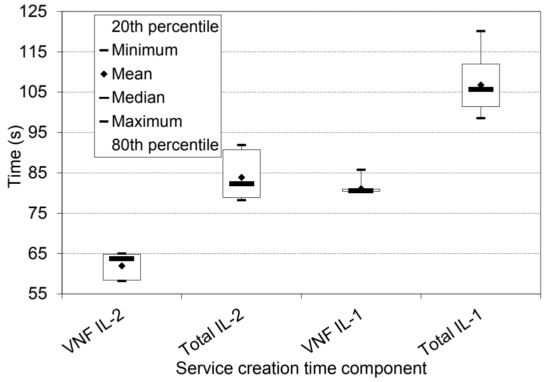
V. SUMMARY AND CONCLUSIONS

An implementation of the 5GT architecture presented in this article has been developed and released as an open-source software implementation [2]. Based on this, a 5GT platform prototype has been deployed in practice running a service for a vertical providing entertainment and media services through a vCDN Network. Our evaluation demonstrates the feasibility of the 5GT platform for automating the instantiation of a complete vCDN service in minutes, as opposed to hours that are the norm in production networks. Moreover, the 5GT system has been tested with 4 additional vertical industries to validate its multi-vertical goals.

As future work, the vertical services automated instantiation achievements obtained with 5GT are expected to be extended to the *operational domain*. To do so, AI-based



(a) Lower values



(b) Higher values

Fig. 4. Boxplot of service creation time components

closed-loop management solutions are envisioned with a special focus on *network management automation* as well as *anomaly detection* for NFV systems.

REFERENCES

[1] A. Oliva et al., "5G-TRANSFORMER: Slicing and Orchestrating Transport Networks for Industry Verticals," *IEEE Communications Magazine*, vol. 56, no. 8, pp. 78 – 84, 2018.

[2] 5G-TRANSFORMER, "5G-T Open Source Software Release," https://github.com/5g-transformer, accessed: 2020-04-01.

[3] Alcardo Alex Barakabitze et al., "5G network slicing using SDN and NFV: A survey of taxonomy, architectures and future challenges," *Computer Networks*, vol. 167, February 2020.

[4] GSMA, "From vertical industry to network slice characteristics," 2018. [Online]. Available: https://www.gsma.com/futurenetworks/wp-content/uploads/2018/09/

[5] M.Chahbar et al., "Towards a Unified Network Slicing Model," in *Conference on Network and Service Management, CNSM*, 2019.

[6] G.Yilma et al., "On the Challenges and KPIs for Benchmarking Open-Source NFV MANO Systems: OSM vs ONAP," April 2019. [Online]. Available: https://arxiv.org/abs/1904.10697

[7] X. Costa-Pérez et al., "5G-Crosshaul: An SDN/NFV Integrated Fronthaul/Backhaul Transport Network Architecture," *IEEE Wireless Communications*, vol. 24, no. 1, pp. 38–45, February 2017.

[8] 5G-TRANSFORMER, "5G-TRANSFORMER final system design and Techno-Economic Analysis," http://5g-transformer.eu/.

[9] 3GPP, "Management and orchestration; 5G Network Resource Model (NRM), TS 28.541, V16.3.0," Dec. 2019.

[10] C. Cassetti et.al., "Arbitration Among Vertical Services," *IEEE PIMRC*, 2018.

[11] K. Antevski et al., "Resource orchestration of 5g transport networks for vertical industries," *IEEE PIMRC*, 2018.

[12] 5G-Transformer, "D5.3: Experimentation Results and Evaluation of Achievements in terms of KPIs."

[13] Josep Mangues, et al., "5G-Crosshaul D5.2: Report on validation and demonstration results," December 2017.

Xi Li is a Senior Researcher on 5G Networks R&D at NEC Laboratories Europe, focused on research of 5G and Beyond Networks. She is the Technical Manager of 5GPPP project 5Growth and has led work packages on the orchestration of 5G mobile networks in the 5GPPP projects 5G-Crosshaul and 5G-TRANSFORMER.

Thomas Deiss is working for Nokia since 1999. He has contributed to standardization on automated testing and worked in requirements engineering for different areas of WCDMA, LTE, and NR base stations, currently on NR radio access. He participated in the 5GPPP projects 5G-Crosshaul and 5G-TRANSFORMER.

Josep Mangues-Bafalluy is Senior Researcher and Head of the Communication Networks Division of the CTTC. He has participated in various roles (incl. leadership) in several public funded and industrial research projects (e.g., 5GPPP 5Growth or 5G-REFINE). He was vice-chair of IEEE WCNC 2018 (Barcelona).

Jorge Baranda Hortigüela is a Senior Researcher in the Mobile Networks Department at CTTC. He has participated in several European, national and industrial projects related with SDN/NFV based orchestration of mobile networks, mobile backhauling and novel wireless communication systems.

Xavier Costa-Perez is Head of 5G Networks R&D and Deputy General Manager at NEC Laboratories Europe. His team contributes to products roadmap evolution as well as to European Commission R&D collaborative projects and received several awards for successful technology transfers.

Giada Landi is an R&D project manager at Nextworks, has participated in many industrial and European research projects. She holds 10+ years experience in telecommunication networks, with focus on control plane architectures and protocols, and consultancies on PCE, SDN, and NFV topics.

Carlos J. Bernardos holds a PhD in Telematics from UC3M, where he works as an associate professor at UC3M. His current research interests are network virtualization and wireless networks. He is an active contributor to the IETF.

Paola Iovanna is a Master Researcher at Ericsson, leading a research team on networking and automation (SDN/NFV) solutions for 5G networks. She is authors of more than 70 patents and 80 publications in either international scientific journals or conferences.

Arturo Zurita is a SDN/NFV engineer at ATOS. His work is on development of SDN/NFV technologies for 5G mobile networks, with a special focus on networking, virtualization and cloud computing technologies.

Philippe Bertin is a Senior Research Engineer at Orange Labs, driving future networks research projects with bcom. His research addresses the design of distributed and dynamic control and data planes for flexible and convergent 5G networks leveraging on SDN and NFV.