A double-tier MEC-NFV Architecture: Design and Optimisation

Vincenzo Sciancalepore, Fabio Giust, Konstantinos Samdanis, Zarrar Yousaf

NEC Laboratories Europe, Germany, E-mail: {name.surname}@neclab.eu

Abstract—The emerging requirements for 5th generation (5G) mobile networks lead to a complete network paradigm refurbishment by leveraging on Network Function Virtualisation (NFV) and Mobile Edge Computing (MEC). While NFV adds flexibility by allowing network functions to be dynamically deployed and inter-connected, MEC brings intelligence at the edge of the mobile network, reduces latency and enhances the offered capacity. This paper analyses the compound effect of simultaneously considering virtual network functions and MEC applications deployed on the same network infrastructure. The proposed architecture aligns and integrates the MEC system with the NFV Management and Orchestration (MANO) by introducing a management subsystem that enriches the MANO with application-oriented orchestration capabilities. This opens new opportunities on jointly managing applications and virtual network functions by allowing MANO-related operations to be triggered by virtual applications runtime processes or MEC services.

I. INTRODUCTION

In deploying the next generation of mobile networks (5G), mobile network operators (MNOs) endeavor in a twofold mission that, on the one hand, is looking at enhancing traditional services (e.g., telephony, web and multimedia), and, on the other hand, aims at integrating in a single network infrastructure new vertical segments for public safety, healthcare, utilities management, connected vehicles and industrial automation [1].

Such technology leap is enabled by the *virtualisation* and *softwarization* of the network infrastructure, which are pushing MNOs towards quicker network upgrades at lower costs, with the objective to build a flexible network infrastructure able to accommodate a plethora of diverse new services. In order to leverage the agile and elastic characteristics of cloud technology, the telco industry is working towards developing systems that will enable the *cloudification* of the existing network architecture and service provisioning. In this context the Network Functions Virtualisation (NFV) [2] is being widely considered as the pillar technology enabler for 5G networks. This disruptive concept changes the way networks are designed, deployed and how services were managed. In particular, NFV allows network functions to be realized over virtual machines, creating a Virtualised Network Function (VNF) instance that is deployed over a Virtualised Infrastructure (VI), such as a datacenter. Multiple VNFs of different

types can be chained to provide a Network Service (NS). Typically, a VI hosts hundreds of thousands of VNFs and NSs, where the fundamental challenge lies in managing and orchestrating the virtualised resources while delivering carrier-grade services. This is the reason why an NFV Management and Orchestration (MANO) framework has been proposed by ETSI NFV Industry Specification Group (ISG), providing facilities for Life Cycle Management of VNFs and NSs.

Framed as an ETSI ISG, Mobile Edge Computing (MEC) focuses on evolving the mobile network's edge, in order to create a cloud-like environment close to the Radio Access Network that hosts enhanced services provided by the MNO or third parties. Such services span across caching for Content Delivery Network (CDN), RAN analytics, vehicular communications, IoT systems, benefiting applications a closer deployment to the User Equipment (UE) [3]. Such Mobile Edge (ME) applications run in form of virtualised objects on top of a generic cloud infrastructure located within the RAN, referred to as the Mobile Edge host. The Mobile Edge management system is responsible for managing both the infrastructure and the ME application instances that run on a single or different Mobile Edge hosts.

Since the MEC management and orchestration system has operating characteristics similar to the NFV MANO, we argue that jointly orchestrating VNFs and MEC applications can provide several benefits from both the infrastructure cost and operation perspective, i.e. CAPEX and OPEX. The former is the ability of using edge cloud platforms commonly to support both applications and virtualised functions. The latter is the need of a common management and orchestration system as an enhanced version of the current NFV MANO, referred to as MANO+ in this work.

In this paper the MANO+ is described in detail elaborating how to attain a joint application and VNF management without incurring in additional hardware and coordination overhead. In addition, MANO+ further introduces flexibility when creating the service chain, as VNFs and ME applications can be combined to address specific service requirements.

The rest of the paper is organized as follow. Section II briefly overviews the NFV and MEC architectures. Section III describes the proposed application-oriented management system, both from a general perspective and from the MEC-viewpoint. Section IV introduces fundamental use cases wherein our system significantly impacts on, and, lastly, Section V concludes the paper.
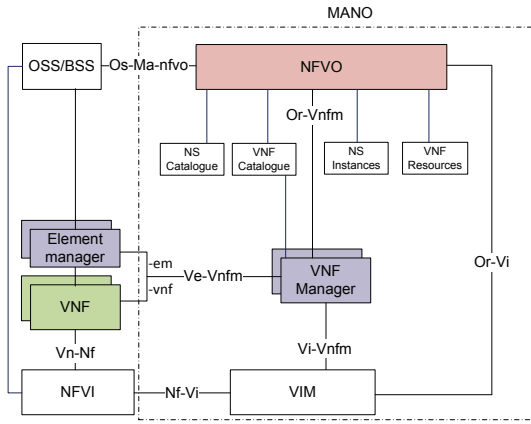
Fig. 1. NFV reference architecture.

Fig. 2. MEC reference architecture (simplified).

II. RELATED WORK

An exhaustive study on network services is presented in [4], considering different deployment environments, and with different optimisation objectives, e.g., resource allocation optimisation, QoS guarantees to mention a few. The application of the SDN paradigm and the NFV architecture to the IEEE Next Generation Service Overlay Network (NGSON) [5], in order to compose a service taking into account contextual information retrieved from specific applications are shown in [6]. The authors of [7] focus on optimising the overall latency in creating service chains of virtual functions considering the data path while connecting them. Lastly, a clearer application-oriented view is proposed by the cross-layer resource orchestration approach in [8].

Differently from the above-mentioned works, in our view we identify MEC and NFV as two complementary technologies to deliver an end-to-end service that consists of *network functions* and *applications* under the control of a unified management and orchestration system. We hereafter provide an overview of the relevant entities composing the two architectures shedding the light on main benefits and limitations.

A. Network Function Virtualisation

The inherent advantages offered by NFV introduce the challenge of the management and orchestration associated with the distributed VNF-Components deployed across multiple servers in a Network Functions Virtualised Infrastructure (NFVI), for providing carrier-grade service. The NFVI consists of the physical resources, such as the computing, storage and network resources that enable the instantiation and execution of VNFs and including the corresponding Element Manager (EM). The lifecycle management of physical and virtual resources is performed by the ETSI ISG NFV proposed NFV MANO framework depicted in Fig. 1. The proposed MANO framework architecture is composed mainly of three functional blocks namely the Virtualised Infrastructure Manager (VIM), the VNF Manager (VNFM) and the NFV Orchestrator (NFVO), interconnected over specific reference points. There are additional data repositories that may contain necessary
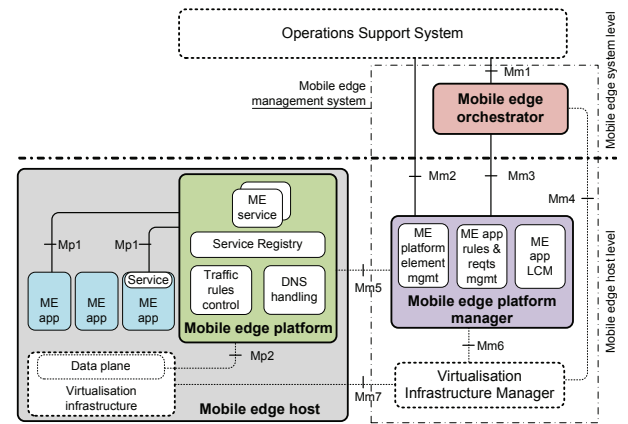
information about NS, VNF, NFV and NFVI that will enable the NFVO to perform its tasks. The MANO architecture also defines reference points for interfacing the MANO system with external entities like NFVI, OSS/BSS, VNFs and EMs for delivering a unified management and orchestration of a VNF system while allowing the NFV to be integrated into an existing network-wide management landscape. The entire NFV system is driven by a set of metadata describing the Service, VNFs and Infrastructure requirements, which are feed to the NFV Management and Orchestration system, in order to act accordingly.

An Interfaces and Architecture (IFA) WG has been formed under the ETSI NFV that develops specifications for the MANO framework. In this respect, the IFA WG at present is in the process of specifying interfaces, requirements and operations for the reference points in view of the functional/operational scope of the NFVO, NFVI and VIM as described in [9]. Besides traditional Fault Configuration Accounting Performance and Security (FCAPS) management, the MANO framework focuses on newer management aspects introduced by NFV such as the creation and life-cycle management (LCM) of the virtualised resources for the VNF, and collectively referred to as VNF management [9]. There are several VNF management tasks such as VNF scaling, migrating, updating, to name a few.

B. Mobile Edge Computing reference architecture

In the MEC reference architecture [10], the Mobile edge (ME) system consists of a set of ME hosts and the associated management entities, see Fig. 2. The ME host is the logical entity that contains the ME platform and the virtualisation infracture on which the ME applications run. The ME platform contains a set of baseline functionalities that enable ME applications to run on a particular host, as well as to discover and consume ME services, or to advertise and provide them through the service registry. The ME platform is also responsible for enforcing the traffic rules to route the data packets to/from the ME applications, as well as to maintain a DNS necessary to discover the ME applications. ME applications

run on the ME host as virtual machines and are designed to consume and/or provide ME services.

The ME management system comprises the ME Orchestrator (MEO), the ME Platform Manager (MEPM) and the virtualisation infrastructure manager (VIM). The MEO has the view the whole ME system, as it maintains the information about all the deployed ME hosts, the services and resources available in each host, the ME applications that are instantiated and the topology of the network. It is also responsible for installing the ME applications in the system, checking their integrity and authentication and validating the associated policies. The MEO has a reference point with the Operations Support System, which is generally the highest level management entity in a mobile network.

The MEPM is associated to a single host, and it is responsible for the ME platform element management, the ME application LCM, and the host level policies management. The MEPM performs such operations with the support of the MEO through a reference point, which is also used to maintain up to date information on available ME services in the ME system. In addition, the MEPM interacts with the OSS for fault configuration and performance management.

The VIM manages the virtualised resources allocated to the ME apps. It has a reference point to the MEO used for management of the application images and the virtualised resources as well as for monitoring the resource availability. Also, it interacts with MEPM to manage the virtualised resources related to the ME applications LCM.

For the sake of brevity, we have omitted some entities from the MEC reference architecture (e.g., the Customer Facing Service portal, the User Application LCM proxy and the associated reference points). Interested readers may refer to [10] for a comprehensive description.

III. COMPOUND ARCHITECTURAL ANALYSIS: MANO+

Given the lack of common orchestration functions for both solutions as well as the number of commonalities that associate NFV with MEC architectures in terms of functional and management descriptions, we pioneer a novel network orchestration architecture. The objective is to cast into a single enhanced MANO platform, dubbed MANO+, both virtualised network functions support and MEC applications with purpose of optimising the service chains creation. To the best of our knowledge, this is the first attempt towards a compound analysis of both solutions by showing design limitations and required amendments to create a solid and practical framework tailored to end-user needs. To this aim, we introduce novel functional concepts and reference points hereafter fully described, into the existing MANO architecture.

A. Virtual Application Functions support and management

An application function is an entity that relies on the functional and/or operational characteristics of VNFs and/or one or multiple NSs to carry out the operations it is designed for. The VNFs may also benefit from the application layer for enhancing operational and/or functional management.
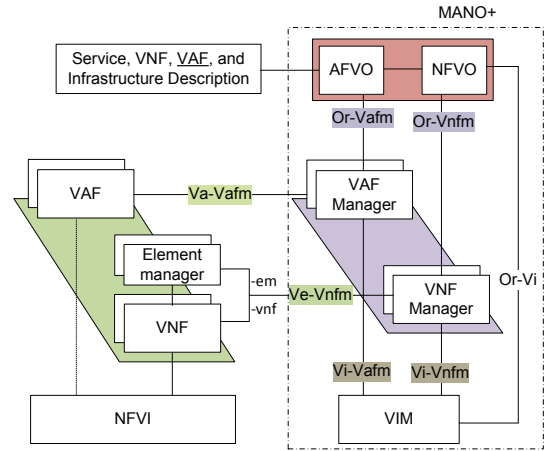

Fig. 3. Extension to the NFV architecture for application oriented management (MANO+).

Interestingly, the two different behaviors devise a service hierarchy wherein the application functions can be added or removed without compromising the functional operations of the underlying VNFs and/or NSs. Although NFVs are used to deploy and manage the underlying NS, it is not as accurate in describing the objects belonging to the upper layer, and their interactions (e.g., through APIs) with the NS.

Let us now consider an example wherein an application function leverages data analytic tools to retrieve results for a given performance metric, by exchanging service primitives with some VNFs through a number of APIs. Although such an application function is not strictly necessary for the specific Network Service lifecycle (actually it can be excluded from the forwarding graph of said NS), obtained results may be used to fine-tune the network service management process. To a broader extent, runtime characteristics of higher layer application functions may feed the optimisation of management procedures in the MANO system.

In order to account for such characteristics, the MANO system should allow for describing managed objects with tools that reflect the application's runtime requirements and interactions. For this purpose, we have identified three logical building blocks to amend the current NFV architecture, with the required interconnection among them (see Fig. 3):

- **Virtual Application Function (VAF)**: it describes an application function based on the requirements of the processes that constitute its runtime (e.g., state machine transitions, API transactions, user session characteristics, network sockets, etc.), and it enables influencing the MANO operations according to those requirements;
- **VAF Manager (VAFM)**: it is in charge for the application functionality management and lifecycle management of a VAF, as well as for relaying the communication between a VAF and the orchestrator (e.g., to convey the runtime requirements mentioned above);
- **Application Functions Virtualisation Orchestrator (AFVO)**: it handles the service orchestration as per the operational requirements demanded by the VAFs that
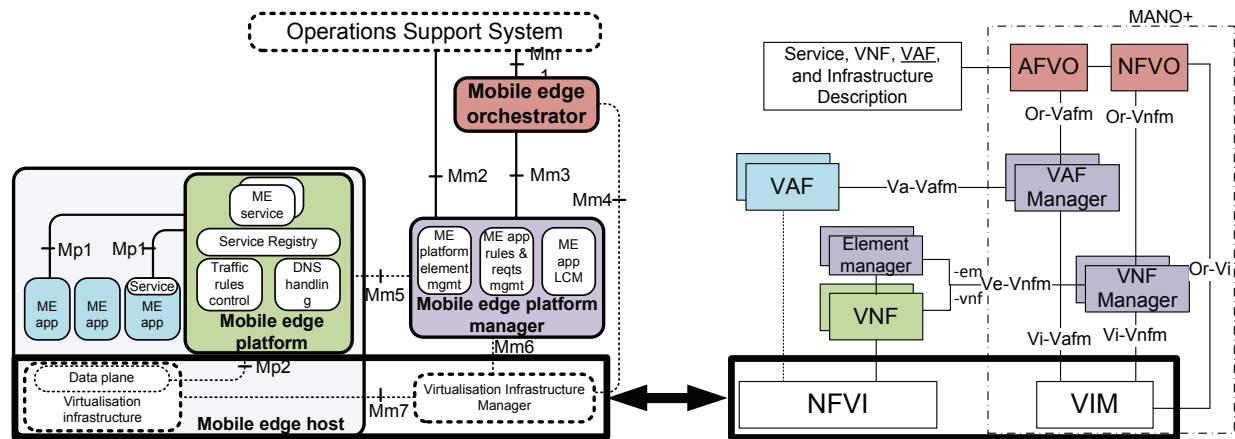
Fig. 4. Compound architectural evaluation of MEC and NFV structural blocks.

support a given end-to-end network service.

The rationale behind the above extension is to tailor the management functions to the application requirements, and at the same time, exposing a common set of interfaces to the management and orchestration system. Therefore, network flexibility is enriched with the application performance management, to optimally orchestrate the NS based on user performance experience. This enables application-driven or even Quality of Experience (QoE)-based MEC service instantiation, which can be easily combined with VNF lifecycle management operations, such as instantiation, scaling, migration or cloning.

The functional scope of VAFM is similar to the MANO VNFM, as it renders VAF lifecycle management, such as VAF instance(s) instantiation/configuration, VAF software upgrade/update/modification, scaling in/out/up/down, healing, instance termination, feasibility/integrity checking, and so on. Additionally, the VAFM manages runtime operations of VAFs, e.g., when the application is distributed among several platforms and needs coordination or for contextual information transfer. Lastly, we envision this entity managing performance measurement results and faults/events information between the VAF(s) and the VNFs in the underlying MEC service platform. Such information is exchanged through the AFVO, which in turns leverages on the NFVOs resource orchestration functions, when necessary to optimise the service.

The new functional blocks require the following reference points:

- **Orchestrator-VAFM (Or-Vafm)** reference point: the collection of interfaces between the MANO Orchestrator and the VAFM;
- **Virtual Application-VAFM (Va-Vafm)** reference point: the collection of interfaces between a VAF and its VAFM;
- **Virtual Infrastructure VAFM (Vi-Vafm)** reference point: the collection of interfaces between a VAFM and the VIM.

Such reference points define several interfaces to support relevant operations, i.e., the performance management interface can prescribe all operations relevant to performance manage-

ment, or other interfaces may support resource management, fault management and policy management. The repositories are extended to support both VAFM and MANO connected to the orchestrator by including *i) VAF Descriptor* (VAFD), a template describing a VAF in terms of deployment and operational behaviour requirements, *ii) VAF Catalogue*, a repository of on-boarded VAF packages including the VAF Descriptor. Based on those requirements, the Orchestrator (NFVO+) will map and connect the VAFs to the appropriate service function(s). The VAFD may also contain application KPIs requirements that can be used by the NFVO+ to ensure that the application function(s) is(are) getting the appropriate level of service from the underling MEC service platform.

Please note that the proposed architecture is a logical description of functionalities that can be either realized by standalone building blocks, or by expanding the existing entities as suggested by the color scheme illustrated in Fig. 3.

B. MEC and NFV: Orchestration and Coordination

The proposed extensions can accommodate the functional blocks that constitute the MEC reference architecture as per the mapping suggested in the following and depicted in Fig. 4. The motivation for having entities from the MEC ecosystem alongside virtual network function is sharing the same virtualisation infrastructure and the related VIM, since both layers provide infrastructure resources, including compute, network, storage, etc. Then, the mobile edge platform should be realized as a VNF, whereas the Mobile edge applications can be mapped to VAFs. With respect to the Mobile edge platform manager, the ME platform element management function can be associated to the VNF EM, whereas the rest of components (ME app rules & requirement management and the ME App LCM), which are related to the Mobile edge application, can be part of the VAFM. Note that the MEC architecture has not defined any entity for the ME platform LCM, but those functions could be easily grouped under the usual VNFM. Last, the Mobile edge orchestrator should devolve the resource orchestration functionalities to the NFVO, and the functions related to application orchestration to the AFVO.

We argue that the proposed extensions as well as the integration with the MEC architecture are essential to reflect the coordination among MEC applications and network functions. Applications contribute to provide and/or enhance a network service (e.g., video throughput optimisation), and they leverage the existing network functions. Our proposal deals with fostering the development and deployment of applications in order to be the major but not the only driver by having applications in the service chain of a network service. A hard dependency could be easily identified between a particular application and one or multiple network functions, when specific performance targets must be achieved. Our proposed framework leverages such flexible service composition to fulfil network service needs. For instance, a MEC application, i.e., a VAF, may require one or more services that are instantiated as VNFs. In this view, the VAF Manager is supposed to collect the information to assess that the underneath network functions are providing the expected performance level. When such requirements are not met, the NFVO+ is triggered in order to apply the appropriate actions to manage the MEC service and VNFs accordingly.

The novel orchestrator includes the logic to build a network service with both network functions and applications. The NFVO+ retrieves the necessary information from the Service, VNF and Infrastructure descriptors, and from VAFs descriptors too. This architecture enables to deploy a MEC management functionality alongside an NFV MANO system, and both management systems would be orchestrated by the same entity, the NFVO+. Such an enhanced orchestrator can be realized as single element, or split into two components: an orchestrator dedicated to MEC, and another dedicated to the NFV system, that can communicate in a peer-to-peer fashion. The latter option allows for flexible and modular deployment of the entities, but it advocates for a standardized interface to grant multi-vendor interoperability. In both options, the network extends its set of network services from a collection of network functions to a collection of network functions and applications (e.g., MEC-based). Intuitively, the extended set of network services might include video and selected traffic optimisation, machine-type-communication, vehicular communications, and so further, as described in what follows.

IV. FEASIBILITY ANALYSIS: USE CASES

In this section, we present three use cases that leverage on the joint MEC-NFV architectural solution to improve their performance . Such user scenarios lay the basis for the novel concept of network multi-tenancy, wherein the MEC concept plays as key-enabler [11]. Please note that foregoing use cases are not intended to be exclusive or exhaustive. However, they provide a solid basis for evaluating and fostering the adoption of such novel MEC-NFV architectural proposal.

A. MEC service-assisted network QoS provision

A Mobile edge platform is designed to offer enhanced services to Mobile edge applications, and the Radio Network Information Service (RNIS) is a key feature within
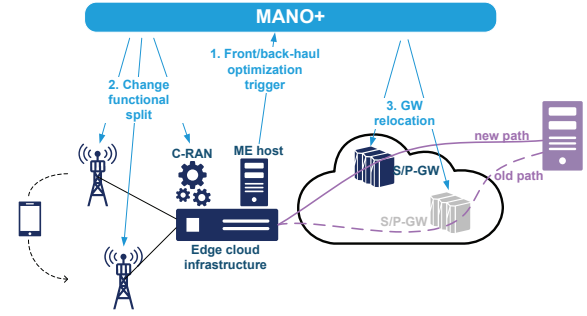

Fig. 5. Operational flow for video streaming applications when MEC-NFV joint orchestration is in place.

MEC. Such service provides an API to ME applications to retrieve relevant information about the radio conditions on different metric basis (per user, group of users and so on). For instance, in a cloud RAN deployment where different functional splits may be applied, RNIS may be exploited by an ME application in order to compute the performance metrics of a given split. In our view, as depicted in Fig. 5, the radio characteristics feed such an application, that, in turn, triggers the MANO+ to improve the performance of attached users considering different functional splits. In particular, the MANO+ may examine alternatives in splitting the base station based on the performance resulting from application changes due to radio conditions or fronthaul dynamics. The MANO+ is required to apply a different functional split following two policies: *i)* increasing the capacity in the fronthaul by shifting RAN functions from the centralized Baseband Unit (BBU) towards the edge or Remote Radio Head (RRH), or *ii)* shifting base station functions from the RRH towards the BBU so as enabling cooperative multi-point (CoMP) schemes or better scheduling and interference coordination. Alternatively, a MEC fronthaul/backhaul optimiser function may trigger the MANO+ to provision changes in virtualised functions of the core network, including for example the re-location of a Serving/PDN-Gateway [12] by shifting a virtual machine into a new location that guarantees a delay reduction or releases resources in the backhaul. This clearly sheds light on the reason why a novel MANO+ architecture is needed.

B. MEC-assisted migration for gaming applications

The advantage of the proposed approach finds a natural application when gaming scenarios are considered [13], as mobile edge applications developed for gaming, virtual and augmented reality, take large benefits from the very low latency and computing offload capabilities offered by the mobile edge host. Nevertheless, the ME application should consider the application runtime state before performing LCM operations. A good example is provided by a gamer moving into an area served by multiple mobile edge hosts. Whereas the mobile edge management system may consider relocating the mobile edge application (or part of its memory state) to better fulfil the latency requirements, the relocation execution may happen in a crucial phase of the game, causing an annoying service degradation as well as a bad user experience.
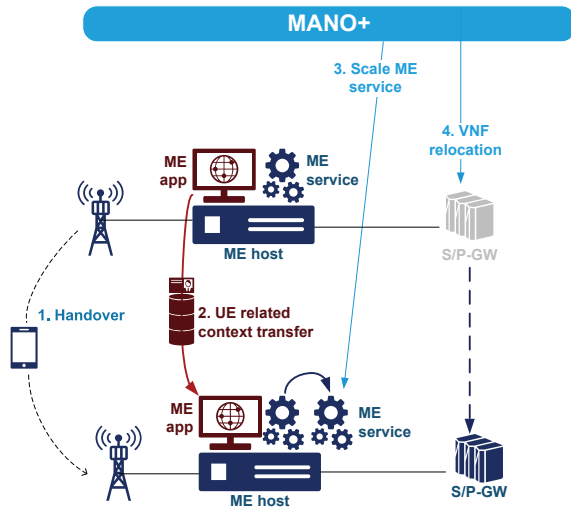
Fig. 6. Service provisioning process triggered by the VAFs mobility.

In our novel architecture, a ME application is instantiated as a VAF, and therefore connected to the VAF orchestrator through the VAF manager. This enables the ME application to influence the LCM operations by, e.g., sending triggers to start operations, or communicating relevant parameters to feed the orchestration algorithm. Therefore, the ME application may suggest a time frame wherein the relocation is recommend to prevent dramatic service degradations, e.g., after the gamer reaches a checkpoint and the game loads the next level.

C. MEC support for user mobility

User mobility issues play a key-role in the mobile cellular networks, as session continuity needs to be guaranteed after handovers and content should be placed efficiently and in such a way to minimize the service degradation [14], [15].

In our proposal, the mobility support is realized by transferring the user-related application content and other contextual information from the application running in the source cell to the application instance running in the target cell under the coordination of the MANO+ system. In fact, when a handover takes place, the role of the joint orchestrator is to make sure that the VAFs service requirements are satisfied in the new location. Hence, MANO+ architecture triggers dynamic adjustments to the VNFs in the target cell to meet the necessary service level, such as additional ME service instances for load balancing and/or relocating of S/P-GWs.

Fig. 6 clearly reflects this use case. A UE is connected to a ME application that depends on a ME service (e.g., RNIS). The ME application and ME service are running on the same ME platform as a VAF and a VNF, respectively. Upon a handover occurs, the UE might move to a cell associated to another ME platform (step 1) causing a coordination of VAF and its VAF manager for the context transfer between the applications running in the source and target cell (step 2). This may require an increasing demand for the RNIS in the target cell, promptly handled by the MANO+ which initiates an additional VNF by scaling out the RNIS component (step 3). While a new ME platform closer to the user can benefit from

the proximity, a pre-assigned S/P-GW may cause backhaul congestion as well as an increasing delay while impairing other user communication performance. Therefore, MANO+ could also perform a S/P-GW re-location to properly optimise the data path for the UE new location (step 4).

V. CONCLUSIONS

In this paper we have evaluated the compound effect of NFV, mostly focused on the infrastructure aspect, and MEC application-oriented, when merged together in a novel proposed framework, termed as MANO+. We have extended the NFV architectural framework by introducing the Virtual Application Function concept intended as a more specific VNF capable of accounting requirements and, at the same time, triggers from the application runtime states and operations.

Our proposal brought new perspectives to the optimisation processes performed by the NFV MANO system, taking into account a new dimension given by the presence of the application layer while being compliant with the last MEC ISG directives. Benefits and limitations of our solution has been extensively evaluated when applied to real use case scenarios.

REFERENCES

[1] NGMN Alliance, "NGMN 5G White Paper," Tech. Rep. Version 1, Februaryl 2015.
[2] ETSI NFV ISG, "Network Functions Virtualisation (NFV); Architectural Framework," ETSI, DGS NFV 002, Decemberl 2014.
[3] ETSI MEC ISG, "Mobile Edge Computing (MEC); Technical requirements," ETSI, DGS MEC 002, April 2016.
[4] W. John, K. Pentikousis, G. Agapiou, E. Jacob, M. Kind, A. Manzalini, F. Risso, D. Staessens, R. Steinert, and C. Meirosu, "Research Directions in Network Service Chaining," in *Future Networks and Services (SDN4FNS), 2013 IEEE SDN for*, Nov 2013, pp. 1–7.
[5] "IEEE Standard for the Functional Architecture of Next Generation Service Overlay Networks," *IEEE Std 1903-2011*, pp. 1–147, Oct 2011.
[6] F. Paganelli, M. Ulema, and B. Martini, "Context-aware service composition and delivery in NGSONs over SDN," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 97–105, Aug 2014.
[7] B. Martini, F. Paganelli, P. Cappanera, S. Turchi, and P. Castoldi, "Latency-aware composition of Virtual Functions in 5G," in *Network Softwarization (NetSoft), 2015 1st IEEE Conference on*, pp. 1–6.
[8] W. Cerroni, M. Gharbaoui, B. Martini, A. Campi, P. Castoldi, and F. Callegati, "Cross-layer resource orchestration for cloud service delivery: A seamless SDN approach," *Computer Networks*, vol. 87, pp. 16 – 32, 2015.
[9] ETSI NFV ISG, "Network Functions Virtualisation (NFV); Management and ORchestration," ETSI, DGS NFV-MAN 001, Decemberl 2014.
[10] ETSI MEC ISG, "Mobile Edge Computing (MEC); Framework and reference architecture," ETSI, DGS MEC 003, April 2016.
[11] K. Samdanis, X. Costa, and V. Sciancalepore, "From Network Sharing to Multi-tenancy: The 5G Network Slice Broker," *IEEE Communications Magazine*, July 2016.
[12] T. Taleb, Y. Hadjadj-Aoul, and K. Samdanis, "Efficient solutions for enhancing data traffic management in 3gpp networks," *IEEE Systems Journal*, vol. 9, pp. 519–528, 2015.
[13] K. T. Chen, Y. C. Chang, H. J. Hsu, D. Y. Chen, C. Y. Huang, and C. H. Hsu, "On the quality of service of cloud gaming systems," *IEEE Transactions on Multimedia*, vol. 16, pp. 480–495, Feb 2014.
[14] D. Munaretto, F. Giust, G. Kunzmann, and M. Zorzi, "Performance analysis of dynamic adaptive video streaming over mobile content delivery networks," in *IEEE International Conference on Communications (ICC)*, June 2014, pp. 1053–1058.
[15] D. Karamshuk, N. Sastry, M. al Bassam, A. Secker, and J. Chandaria, "Take-away TV: Recharging Work Commutes with Greedy and Predictive Preloading of TV Content," *IEEE Journal on Selected Areas in Communications (JSAC)*, Jun 2016.