ETSI White Paper No. 54

Evolving NFV towards the next decade

1st edition – May 2023

Authors (alphabetical order):

Xuhui CAI, Hui DENG, Lingli DENG, Ahmed ELSAWAF, Shen GAO, Arturo MARTIN DE NICOLAS, Yoshihiro NAKAJIMA, Janusz PIECZERAK, Joan TRIAY, Xuliang WANG, Baoguo XIE, Hammad ZAFAR

ETSI
06921 Sophia Antipolis CEDEX, France
Tel +33 4 92 94 42 00
info@etsi.org
www.etsi.org

About the author

This White Paper has been produced as a collective effort within the ETSI ISG NFV, and on its behalf the following editing team (listed in alphabetical order):

Xuhui CAI (China Mobile)

Hui DENG (Huawei, ETSI ISG NFV Vice-Chair)

Lingli DENG (China Mobile)

Ahmed ELSAWAF (Nokia)

Shen GAO (China Telecom)

Arturo MARTIN DE NICOLAS (Ericsson)

Yoshihiro NAKAJIMA (NTT DOCOMO, ETSI ISG NFV Chair)

Janusz PIECZERAK (Orange, ETSI ISG NFV Vice-Chair)

Joan TRIAY (NTT DOCOMO)

Xuliang WANG (China Telecom)

Baoguo XIE (ZTE)

Hammad ZAFAR (NEC)

Contents

Executive Summary

Network Functions Virtualization (NFV) has been the catalyst of a radical change to the telecom industry, leading the transition from the traditional physical (hardware) network appliance into a new software-based virtualized network function (VNF) era. Based on the use of general-purpose (commodity-off-the-shelf, COTS) servers and the deployment of network functions as software applications, NFV has broken through the technical challenges of software and hardware decoupling. The migration of network functions from dedicated physical appliances to distributed cloud infrastructure has revolutionized the way communications networks are developed, deployed, and operated nowadays.

In 2012 various international leading telecom service providers (network operators) jointly released the seminal White Paper about NFV's concept and vision and announced the beginning of a new era in the telecom industry. Since the creation of the ETSI NFV Industry Specification Group (ISG) in late 2012 (the first standards organization of this kind in this domain) network operators, communications technology (CT) vendors, information technology (IT) vendors, small and medium-sized enterprises (SME) and other core contributors (e.g., from open source, academia and research communities) have been actively discussing and standardizing the NFV framework, which has become the telco cloud & virtualization network architecture of reference.

The mixture of IT and traditional telecom networks viewpoints in NFV have also brought a challenging, yet an important network transformation environment. Traditionally, network operators and CT vendors have become accustomed to the consensus-based standard development process, from which products are then developed and commercialized. However, the scenario of IT vendors and open source communities is different; these are based on a "code-first" process, whereby code is first developed and then contributed. These two ways of developing current telecom networks technology need to be brought together, as "sides of the same coin"; in this scenario, network operators, CT and IT vendors, and open source communities need to work together to bridge the gap between the two perspectives and facilitate broader support of NFV standards. For instance, while open source could further support and implement the standards as much as possible, standards could consider the "borrow-in philosophy" to take advantage of the core strength of open source in the specifications.

Ten years after the emergence of NFV, and after major investments and network deployments based on it, it is the right time to consider how NFV can and will evolve. For this, it is important to understand and take actions to the requirements that telecom service providers consider for their telco cloud: 1) unified network management; 2) use of latest cloud-native, IT, automation and Artificial Intelligence (AI) open source software, and 3) multi-vendor interoperability and migration among different clouds.

Building on previous achievements, this White Paper analyzes different challenges and technology trends, and proposes several potential directions on how NFV can evolve in the next decade. Aspects about API development, open source, NFV multi-cloud, unified management, automation, and AI are considered as key drivers for the evolution.

The White Paper is organized as follows: section 1 summarizes NFV's history; section 2 recaps learnings, experiences and challenges related to NFV and introduces various trends and opportunities in network transformation that can play an important role for NFV. Section 3 develops the way forward for evolving NFV. Finally, section 4 concludes the White Paper.

1 NFV: a short history

The NFV concept was first proposed in 2012, almost at the same time as the ETSI ISG NFV was established. The group started its journey from a joint effort among telecom service providers and vendors. Key experts were entrusted to the development of a new approach to provide telecom networks based on virtualized technologies.

ETSI ISG NFV attracted not only vendors of telecom network systems, but also vendors of infrastructure and software enabling virtualization. From the very beginning, it was identified that a joint effort was needed to define requirements for an infrastructure required to be efficient, reliable, and capable (e.g., with acceleration support) in order to build a solid base for transitioning telecom networks towards clouds and NFV. NFV's success highly relies on this collaboration between network operators and vendors, and the consensus reached between them.

By adopting cutting-edge IT technologies, NFV has been increasingly influenced by open source communities, especially considering commercially available network deployments wherein integrated solutions have been built based on key reference open source projects. Indeed, at very early stages it was identified that, in order to realize the NFV from a simple theoretical concept to real commercialization, the ETSI ISG NFV needed to join efforts with open source communities as well, to get benefits from both formal standards and de-facto standards and accelerate the development based on readily available open source based solutions. Examples of such collaboration came in early during the first years of the ISG, like the creation of the OPNFV which was tasked to release NFV integration solutions based on existing IT and open source software. Later, as the network and service orchestration became particularly important for service providers, a number of other open source communities such as OSM, OPEN-O, and ONAP built upon the concepts from NFV.

Initial stages of the ISG focused on checking the feasibility of NFV, which subsequently transitioned into formal development of standards for interoperability. Since then, the ETSI ISG NFV has published dozens of NFV specifications. These have been grouped, released, and evolved through a sequence of "NFV Releases", in a short recapitulation: NFV Release 1 laid the concepts foundation, NFV Release 2 provided the first set of interoperable implementation solutions, and NFV Release 3 enhanced the NFV framework by addressing operational matters. From NFV Release 4, additional support for containerized VNFs towards more cloud-native deployment and automation has been specified. Finally, the ongoing NFV Release 5 is dealing with the consolidation of various ecosystems which have had a profound impact on many NFV related standards organizations and open source communities, including 3GPP, TM Forum, LFN (e.g., Anuket, ONAP, Nephio, etc.).

As a promoter of network transformation and a key enabler of technology evolution, NFV technologies are continuously evolving, starting from the earliest LTE’s core network virtualization based on virtual machine (VM) virtualization technology, to 5G network cloud deployments based on OS container virtualization technology and with basic level of automation capabilities. As a matter of fact, NFV has played a key role in defining 5G. As an example, the 3GPP defined 5G system architecture and the design of 5G core network service-based architecture (SBA) were heavily influenced by NFV, even to a point that one could claim that “5G is an NFV-native network”. Today, practically all product portfolios for 5G core network deployments offered by vendors comprise of virtualized network functions based on ETSI NFV standards.

Recently, with the introduction of the concept of autonomous network, the NFV standards are not scoped to network virtualization matters only (like VM or OS container-based virtualization) but are rather extending as well to other more transversal and multi-faceted technology areas. Examples are the efforts to consider AI for IT operations (AIOps) related services, models and interfaces into scope, such as intent-driven network service (NS) lifecycle management, management data analysis, Platform-as-a-Service (PaaS) and related VNF generic OAM functions.

While the level of achievements and real benefits of NFV might not equate among all service providers worldwide, partly also due to the particular use cases and contexts where these operate, the reality shows us that, based on the ETSI NFV architecture, service providers have been able to build ultra-large-scale telco cloud infrastructures based on cross-layer and multi-vendor interoperability. For example, one of the world's largest telco clouds based on the ETSI NFV standard architecture includes distributed infrastructure of multiple centralized regions and hundreds of edge data centers, with a total of more than 100,000 servers. In addition, some network operators have also achieved very high ratios of virtualization (i.e., amount of virtualized network functions compared to legacy ATCA-based network elements) in their targeted network systems, e.g., above 70% in the case of 4G and 5G core network systems. In addition, ETSI NFV standards are continuously providing essential value for wider-scale multi-vendor interoperability, also into the hyperscaler ecosystem as exemplified by recent announcements on offering support for ETSI NFV specifications in offered telco network management service solutions.

2 Challenges, trends, and opportunities for NFV

2.1 Declarative intent-driven network operations

It is becoming apparent that there is growing complexity inherent to disaggregated network deployments, due to the introduction of virtualization and cloudification technologies (from a vertical network disaggregation perspective), as well as too fined-grained customization experienced in actual network development.

To cope with the growing complexity and to simplify network operations, more generic management solutions are being envisioned and becoming a reality. This can help service providers cover a broader spectrum of use cases, both at the network deployment level, as well as at the operations level with a common generic management toolset. In both levels, NFV technologies bring major opportunities, not only in offering a framework for deploying networks on a more unified physical and virtual infrastructure, but also in providing a generic management and resource services platform. Moreover, simplification drives towards network operations' efficiency and agility, which applies to the network itself and the services offered over the network. This improves customer satisfaction and the perception of telecom operators.

In particular, the introduction of declarative intent-based operations can play a key role to simplify NFV network operations. The declarative intent-driven operation only needs API to be consumed and exposed by/to managed objects' desired state to maintain. This shifts more responsibility to the API producer management function into fulfilling the desired state, as well as it demands additional orchestration capabilities to be supported by it. This more declarative operation manner has been successful in diverse management and orchestration systems, like those dedicated to managing OS containers from an infrastructure point of view, as well as others in the services level management.

2.2 The rise of containerization and heterogeneous infrastructure

When NFV started, virtual machine was the dominant and most mature virtualization technology in the industry at the time. Therefore, it was a natural choice for the NFV community to design the architecture, models, and interfaces around VM-based virtualization technology like those offered by KVM as hypervisor layer, Open vSwitch for virtual networking, OpenStack as virtual infrastructure management, etc. However, NFV has, since the very beginning, the ambition to create a virtualization technology agnostic reference architecture and associated solutions; as examples, concepts such as "virtualization container" and virtualization deployment unit (VDU) were created to abstract the differences of the underlying infrastructure. However, abstracting different virtualization technologies is not an easy task as available upstream solutions to build NFV have very specific technology-dependent factors.

The benefits of VM-based virtualization technology, such as their more dynamic manageability with respect to legacy hardware-based solutions, as well as the good levels of isolation, security and performance have proved to be a key asset in past and current NFV deployments. It could even be said that some network functions benefit greatly by the stability and maturity of the VM-based virtualization technologies, which is a perfect match when considering the stringent long-term support and maintenance requirements of telecom networks.

The OS container virtualization technology (or more commonly referred as containers) is becoming over time the main choice by many developments of cloud-native based micro services applications. Among its advantages are the deployment speed and much smaller footprint, factors that can help in improving the resource utilization and lowering the resources consumption.

Whilst NFV is rapidly developed and deployed in the world, virtualization technologies are also evolving quickly. The virtualization technology represented by micro VM, Kata Container and unikernel combines the advantages of VMs and containers, which not only satisfies the good levels of isolation, security, and performance, but also has other characteristics such as flexibility, rapid deployment, and efficient resource utilization. In addition, WebAssembly (Wasm) is more portable and flexible than container and can also be a good choice as a virtualization technology for edge devices.

As more network operators and vendors are already leveraging the potential of OS container virtualization (containers) technologies for deploying telecom networks, the ETSI ISG NFV also studied how to enhance its specifications to support this trend. During this work, the community has found ways to reuse the VNF modeling and existing NFV management and orchestration (NFV-MANO) interfaces to address both OS container and VM virtualization technologies, hence ensuring that the VNF modeling embraces the cloud-native network function (CNF) concepts, which is a term commonly referred in the industry nowadays. This has been achieved despite OS container and VM technologies having somewhat different management logic and resource descriptions. However, diverse and quickly changing open source solutions make it hard to define unified and standardized specifications. Nevertheless, due to the fact that both kind of virtualization technologies can and will still play a major role in the future to fulfill the various and broad set of telecom network use cases, efforts to further evolve them as well as to complement them with other newer virtualization technologies (e.g., unikernels) are needed.

Furthermore, driven by new application scenarios and different workload requirements (e.g., video, Cloud RAN, etc.), new requirements for deploying diversified heterogeneous hardware resources in the NFV system are becoming a reality. For example, to meet high-performance VNFs, requirements for heterogeneous acceleration hardware resources such as DPUs, GPUs, NPU, FPGAs, and AI ASIC are being brought forward. In another example, to meet the ubiquitous deployment of edge devices in the future, other types of heterogeneous hardware resources, such as integrated edge devices and specialized access devices, are also starting to be considered.

In high-performance, high-integration, and low-cost scenarios such as AI training, video processing, and network security, heterogeneous hardware such as DPUs, GPUs, and NPU can be used to flexibly offload basic service and application workloads, such as big data analysis, network, storage, and security data processing. Heterogeneous hardware such as SmartNICs and smart cloud cards can be used to offload virtual switch flow tables or hypervisors, which can greatly improve network performance such as packet forwarding, storage acceleration and security encryption.

Furthermore, current operator needs extend beyond performance, scalability, and flexibility, into other aspects such as privacy and security. In that realm, the concept of secure isolation and privacy protection are gaining traction. These new infrastructure options cannot simply be considered as another embodiment of existing hypervisor-based or OS-based NFVI due to the introduction of some fundamental elements, such as Security Monitor, and concepts such as encrypted memory space, trust status verification, hypervisor partitioning, etc. Going forward, such security-related elements need to be considered as an additional key building block for addressing new challenging security requirements in NFV.

2.3 Autonomous networking, automation, and unified/sole data source

The autonomous network technology is a mainstream transformation trend in the telecom industry. It aims to enhance the automatic and intelligent operation, administration, and maintenance (OAM) capability in the whole lifecycle of the network, provide end consumers and vertical industry customers a "zero waiting, zero fault, and zero touch" experience, support customer, individual, enterprise and new vertical service development, and build "self-configuration, self-healing, and self-optimization" capabilities for intelligent network OAM. Many different standards organizations, such as ETSI, TMF, 3GPP, GSMA, NGMN, and open source communities, such as the Linux Foundation, have embraced within their respective technical working scopes this technology and formulating standards to jointly promote the development of automation for the telecom ecosystem.

The main objectives of the autonomous network are diverse. The first is reducing operational costs and errors caused from manual operations, and the complexity of system management and maintenance by introducing more close loop automation. Building and maintaining an NFV-based network deployment can be rather complex as the network gets more disaggregated. Various and different components such as resource orchestration and management functions, infrastructure platform and management, the VNFs themselves, etc. need to be all considered. A well-designed autonomous system could help hide the complexity of the interworking between these components, and thus ease the work of human operators.

The second objective is improving the efficiency in developing and helping to reduce the time in launching new services. Here, an example can be the automation capabilities of a system that runs through and automates at several steps in the process such as software version management, integration, testing, distribution, service deployment, and upgrade processes, etc. Figure 1 below illustrates the trend whereby a transition occurs between a case (shown on the left side) where full automation takes places at selected steps in the lifecycle of current telecom networks to a case (shown on the right side) where full automation can take place at each and every step, thus realizing a full end-to-end automated lifecycle.

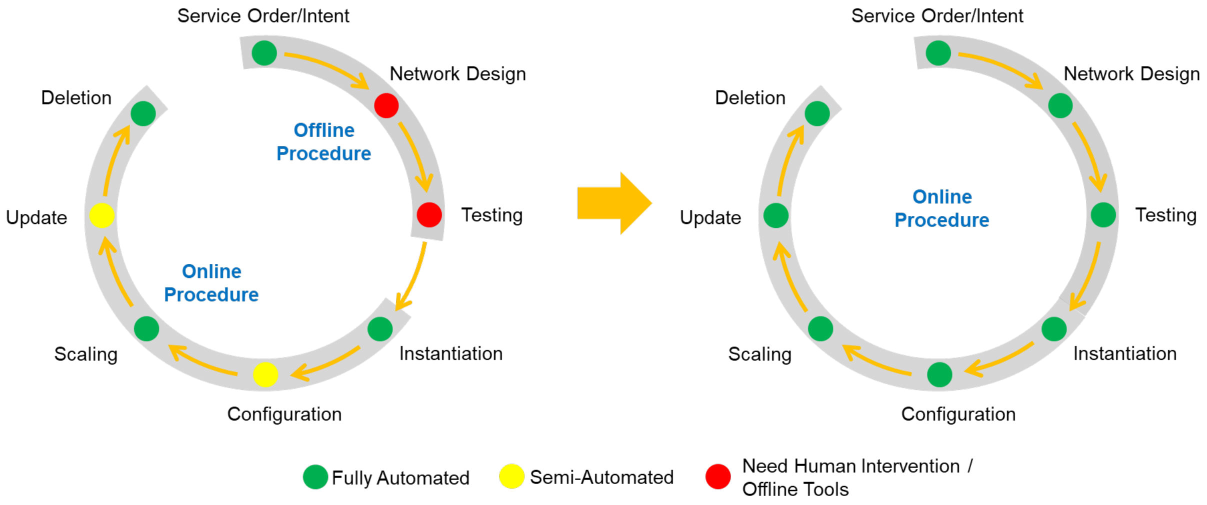

Figure 1: Transition to end-to-end automation

Thirdly, AI, machine learning, and other technologies are being introduced to help network operations personnel handle fault alarms and optimize system performance more effectively by analyzing mass data and building models.

By using the NFV technology, telco cloud management has implemented basic automation of network function deployment, monitoring, and management. However, the trend shows that telco cloud management is evolving towards higher-order self-intelligent networking, by using cutting-edge technologies such as big data, AI, and declarative APIs. Self-service, self-distribution, and self-assured telecom network infrastructure to support telecom networks can open up the door to multi-service, multi-field, full-scenario, and full-lifecycle closed-loop self-intelligence capabilities and provide efficient methods for operators in network planning and OAM.

One basis of implementing automation is data management. Without accurate, real-time, and reliable data, it becomes impossible to drive the system to make decisions and take appropriate actions, thereby hampering the effect of the management system. The current multi-vendor, multi-component and layered construction of an NFV-based network brings additional challenges to data management. The first challenge is data dispersion and non-standardized data: different components in the system collect and maintain their own data for management and orchestration. Furthermore, solutions from different vendors or open source projects usually have different designs with regards to data collection frequency, data format, etc., which makes it harder for data comparison, validation, synchronization and analysis. The second challenge is the limited scope of data collection: to enable end-to-end data analysis, information on all assets in the system need to be obtained; however, current NFV standards and solutions mainly focus on the performance and fault events from the virtualization layer (e.g., VM, containers) and lack necessary information of the underlying hardware.

The third challenge is related to data usage constraints and policies that may need to be considered by the NFV system for management and intent-related data analytics. For example, multiple tenants using the same NFV environment and centralized data analytics functions may have different policies and constraints on data collection, formatting, processing, etc.

In this domain, there are trends in the industry that show new opportunities to be considered. For instance, typically public Cloud/IT systems usually contain a configuration management database (CMDB), which can gather information of all the software and hardware assets in the system, and supports data synchronization, verification, data discovery and retrieval functionalities. In this domain, data meshing, either centralized or distributed depending on the use cases, is becoming prominent as a way to delineate the source of truth of all data sources from workload, infrastructure, user, etc.

2.4 Fragmentation of telco cloud implementations

The successful experience of telecom network operators in the last decades is based on following international telecommunication standards specifying the expected multi-vendor interoperability. On the contrary, the IT industry is more software driven and it relies on its own software development capabilities (mostly based on open source) and de-facto standards to build an ecosystem which help for the integration of multi-vendor solutions.

It can be realized that most of the IT companies have made their own extensions based on the common open source container infrastructure orchestration systems like Kubernetes® (K8s®), and utilizing multiple open source projects from the Cloud Native Computing Foundation (CNCF®), so even though a common K8s® source code solution is leveraged, this does not mean that interoperability with others is ensured, typically due to providing different container-as-a-service (CaaS) functionality and APIs. In such a scenario, an application developer might not be able to successfully deploy its application on different IT CaaS solutions unless corresponding customization is developed.

The ETSI NFV architectural framework inherited concepts both from more traditional layered telecom management frameworks, such as the ITU-T's Telecommunication Management Network (TMN), as well as from cloud management frameworks, which in turn leverage open source solutions. Therefore, it can be said that the NFV architecture adds the necessary support for carrier-grade and telecom requirements on top of leveraged open source frameworks, and it builds the cornerstone for a multi-vendor three-layered NFV system comprised of infrastructure, VNF and network orchestration.

In the last years, CNCF's K8s® has attracted a lot of attention. As part of the efforts by the ETSI ISG NFV to better support the deployment and management of OS container-based VNF, additional NFV-MANO functions have been specified, such as the container infrastructure services management (CISM) and CIS cluster management (CCM). The corresponding NFV-MANO architectural framework specified in ETSI GS NFV 006 has been updated to reflect such additions. The additional functionality is represented in figure 2 as an additional layer, the CaaS layer, as way to depict the groupings of functionality that can happen in an integration of the NFV system when considering container-based deployments. Figure 2 provides a high-level illustration about the major groupings of functionality in an NFV environment, represented as layers; however, it is not meant to represent neither a strict boundary of interfaces interoperability nor a strict delineation of interactions between components (e.g., orchestration layer components can interact directly with underlying infrastructure layer components).

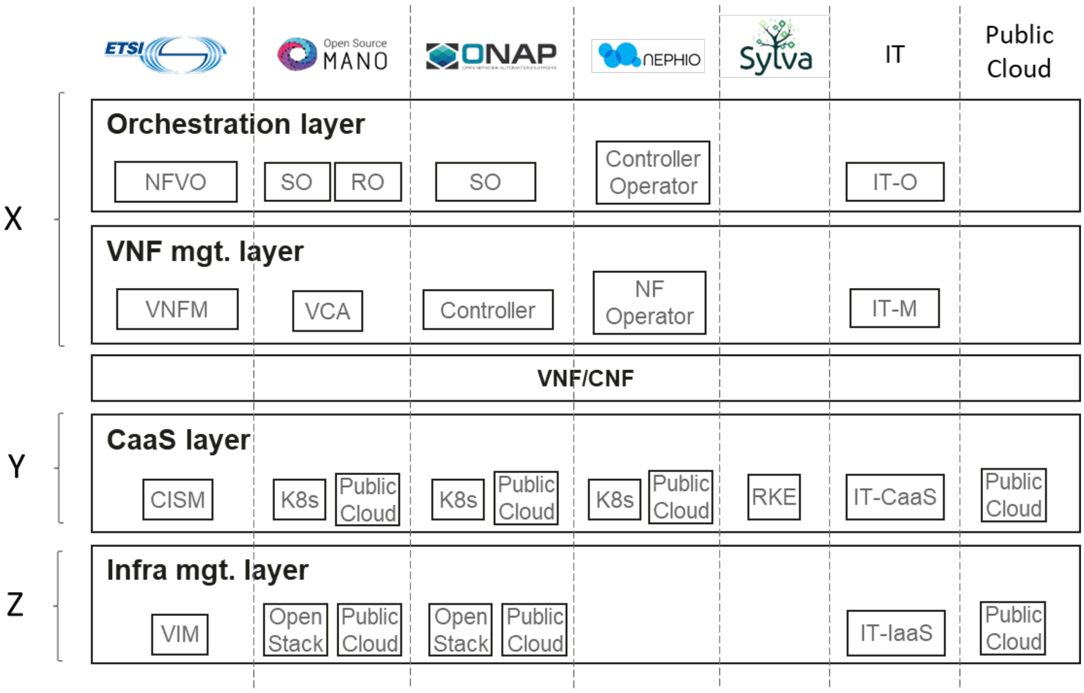

Figure 2: The choices of cloud native implementations for telecom network operators

As shown in figure 2, there are diverse solutions that could eventually be leveraged to realize a cloud native telco cloud:

1. ETSI NFV standard based solutions, covering all the way from the orchestration to the infrastructure layer,
2. Telecom focused open source-based solutions focusing on VNF management and orchestration layers, such as OSM, ONAP, Nephio, etc. which in turn might leverage independently developed lower layer CaaS and infrastructure platforms,
3. Specific container platform solutions, e.g., as projected by the open source project like Sylva,
4. IT based solutions, typically encompassing all identified layers (from orchestration down to infrastructure), and
5. Public Cloud solutions, typically focusing on the lower CaaS and infrastructure management layers.

Because there is no dominant IT solution or telecom open source project for the management and orchestration systems, it is unlikely to have one unified solution solely based on these two (i.e., IT and open source) for many different network operators. In such cases, vendors' VNF/CNFs have to be customized in order to be integrated with those silo solutions, making it harder to truly decouple infrastructure and VNF/CNF. This also brings a high risk of vendor lock-in, in case the network operator wishes to expand management coverage to other network domains or even migrate to some other management system solutions at a later point in time.

As shown in figure 2, for each layer of the telco cloud architecture (network service orchestration, VNF management, CaaS and infrastructure management), there are many choices. As a simple calculation, assuming there are X options for the combination of orchestration and VNF management layers, Y options for the implementation of CaaS layer, and Z options for the underneath infrastructure management layer, to cover the case of VM-based VNF deployments, the number of options of functional integration can be X*Z, but when considering the case of OS container-based VNF deployments, the number of possible combinations becomes X*Y*Z which brings more challenges in deployment, lifecycle management and network operations.

Having a greater number of possible combinations is a positive signal showing the interest on NFV-related technologies, as long as proper interoperability among them can be ensured. However, it is perceived that this is not fully the case, thereby siloed solutions arise which impede the adoption of NFV, complicate their integration, and fragment the whole NFV telco cloud ecosystem in the end.

2.5 Business sustainability versus rapid release evolution of open source

While standards focus on architecture design and interoperability, open source projects prioritize code development. Such rapid development is also the basis for quickly delivering Proof of Concepts (PoC) to implement use cases specified by the standards. However, when it comes to commercialization, integrating the open source code into actual products is not an "overnight" process. Two main challenges are identified on directly using, with no optimization and customization, open source code into telco cloud.

Firstly, on the one hand, the version of open source code and API is updated and iterated rapidly, and some APIs are replaced frequently. For example, K8s® now has a new release every four months. In addition, deprecation of functionalities and APIs can happen systematically. Long-term support is typically not guaranteed by many open source projects. On the other hand, network operators need to provide long-term service and operations for their networks, with many network functions often running for more than five years, or even longer. In such a context, for the sake of business sustainability, the lack of long-term support of open source components brings unexplored challenges in the operations and maintenance process of the network, if open source solutions are leveraged.

Secondly, the telecom network requirements, e.g., delivering a carrier-grade solution, are a challenge for a vast majority of open source projects. For example, the support of CPU isolation, NUMA affinity, and huge page design features, which are very much needed for supporting the deployment of containerized VNFs, require additional development effort, which might not be of interest if an open source based system has been originally envisioned to support other types of workloads, e.g., IT applications. Other examples, like VNF data forwarding and isolation capabilities, demand container networks to support additional features such as multiple interfaces for a Pod, SR-IOV, and DPDK. In summary, mainstream open source projects are often based on the needs of the IT industry, and it is thus a challenge for the network operator and CT vendors to directly plug-and-play the open source code into an NFV-based telco cloud environment.

2.6 Hyper-distributed & full-interconnected edge deployments

As service provider's network deployments extend their coverage, providing more connectivity and services at the edge has become increasingly important. For instance, ultra-reliable and low latency communications (URLLC) demand telecom and compute services to be placed closer to the end user.

Initial deployments of network virtualization, and of NFV in particular, have taken at earlier stages a more "conservative" approach by focusing on deployments of part of the network that are controlled and managed in a more centralized manner, e.g., the core network part of a mobile network. However, as service providers are getting more experienced with NFV, it becomes clear that network virtualization is rapidly extending beyond core network use cases to the edge and access networks. The approach is not strange at all to service providers, since from the very beginning edge computing concepts and developed solutions have also been grounded on the NFV concepts. In addition, virtualization of access networks, such as radio access network (RAN) is becoming a hot topic by multiple organizations, including the ETSI ISG NFV.

With the in-depth digital transformation of industries, edge computing moves down from the convergence edge to the access edge, ultra-edge, or even terminal edge to ensure real-time applications such as industrial devices control and metaverse. Largely speaking, edge computing is based on similar virtualization infrastructure as other parts of the network, when considering NFV. Correspondingly, infrastructure resources expand from centralized deployment to a full and ubiquitous distributed network functions deployment considering various cloud locations such as central, regional and edge, even in some cases to possibly extreme and far reach locations such as air and space.

This trend shows that as network deployments expand, use of NFV-based technologies will also be expanded to cover all network domains. This will bring new requirements regarding the distribution of infrastructure (both physical and virtual), network connectivity, telecom networking and management and orchestration to reach such a highly distributed infrastructure.

3 Towards the next decade of NFV

3.1 Context for NFV: 5G advanced, 6G and beyond

With the gradual adoption of the NFV technology, the wide application of 4G and 5G networks, and the trend of edge computing, the scale of the NFV systems is gradually expanding, and more and more services are being deployed on top of these systems, proving how NFV's telecom cloud infrastructure becomes the heart of 5G networks.

Broad acceptance of NFV in telecom networks sets high demands and expectations towards NFV regarding future evolution of telecom networks. NFV development has proven so far that the use of technological innovations is broadly and effectively adopted. Fast inclusion of necessary enhancements to enable use of NFV for new types of networks (e.g. 5G, RAN) or new capabilities (e.g., network slicing) has shown the flexibility of the technology. Telecom network's transformation is accelerating and NFV needs to respond to this progress by considering new types of infrastructure, virtualization technologies that can accommodate multiple workload types, use cases such as Metaverse and XR, telecom PaaS and SaaS. This will keep NFV attractive for new types of telecom networks (from core to edge) evolving towards 5G advanced, 6G and beyond to further fulfill NFV initial promises.

This White Paper identifies the following areas as key pillars for evolving NFV towards the next decade: API development and further leverage of open source, the creation of a truly multi-cloud and multi-technology NFV environment coping with heterogeneous infrastructure supported by a unified NFV management and orchestration framework enriched by enhanced automation and AI-based technology to achieve real cloud native autonomous networks.

3.2 APIs development and relationship to open source

NFV solutions, and in particular the architectural framework specified by the ETSI ISG NFV, have consciously rethought the approach to perform network operations. The concepts of descriptors or templates used by various management and orchestration procedures, such as the ones defining a VNF or an NS were from the very beginning part of the specification roadmap. These kinds of templates could be seen as the precursor of "model-driven management", which have become later a foundation for developing more advanced "declarative-based management operations".

The original NFV framework has mixed the use of templates with a more traditional imperative operation, which emphasizes the signaling interaction on the interface between involved management functions. A consequence of such an approach is that API consumers need to integrate a lot more orchestration and protocol machinery to fulfill certain network management procedures.

In the next decade, simplification of NFV systems will be further pursued. On the one hand, this will entail simplifying the interfaces of the existing NFV systems by introducing more intent-driven concepts and further adopting declarative API design into the NFV standards in additional management layers. The expectation is that, by jointly considering other end-to-end automation capabilities, this will further reduce the complexity of multi-vendor integration and reduce the workload of development and testing as a means to facilitate interoperability. And this might ultimately imply the need to design and create new APIs and re-assign the responsibilities between the API consumers and its producers.

As an example of this vision, figure 3 illustrates a simple scenario of how declarative intent-driven API will play a key role at the border between network operators' supporting systems and the NFV-MANO system. Intent language could help the NFV-MANO consumer (e.g., network operator) better describe what objectives they want to achieve and hide the complexity of fulfilling the network operations implementation to the NFV-MANO system, and also make it truly feasible for multi-vendor interoperability since network service intents can then be translated into corresponding NF intents and further interpreted into the appropriate lifecycle management procedures for the respective VNFs (and thus also of CNFs).

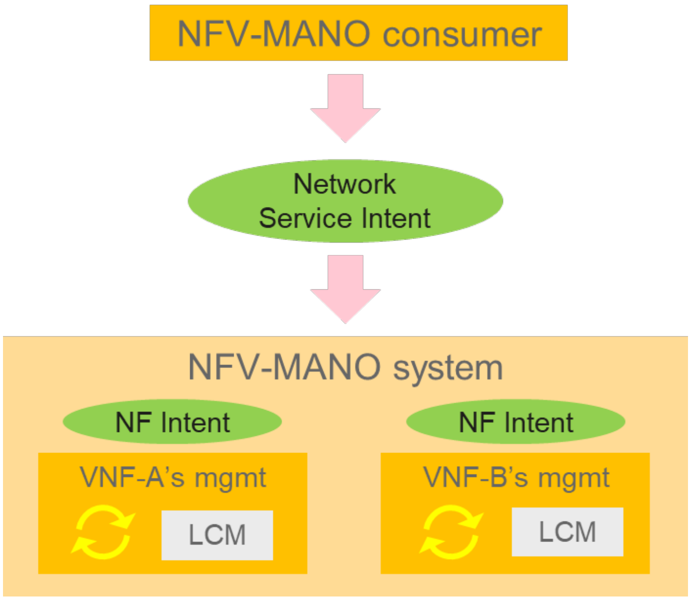

Figure 3: Exemplary simplification of a declarative intent-driven NFV-MANO system

In relation to the importance of well-defined APIs to facilitate interoperability, leveraging emerging IT open source projects (with or without relation to telco cloud), will continue to set the basis for developing the NFV standards. The broad deployment of the telco cloud builds upon two cornerstones (standards and open source) to reach success. Facing the future NFV evolution, ETSI ISG NFV needs to continue expanding the cooperation between both communities to promote efficient open source integration and telecom enhancements. This is expected to help NFV take advantage of the latest technology advancements in the industry.

On the basis of keeping the architecture simple and interoperable, the "borrow-in philosophy" in NFV standards speeds up the development of standards, improves the efficiency of protocol-to-code conversion and avoids standards "over-specification" problem. This will make NFV standards flexible to new technologies and concepts while maintaining their backward compatibility.

NFV standards will need to take particular attention with the rapid development of the open source technologies to excerpting stable versions of APIs provided by the referenced open source solutions (such as K8s®) and avoid relatively immature APIs (such as alpha version APIs) to keep the standard interfaces relatively stable for long-term support and interoperability. Open source organizations like CNCF have provided diverse solutions which could be combined to set a good basis for cloud-native infrastructure, management and orchestration. Together with NFV's focus on the special scenarios and requirements of the telecom network, profiled de-facto APIs can be efficiently leveraged and their usages be proofed in a more complex network and management scenario. Several aspects of maturity, performance reliability, security and industry acceptance need to be considered in the profiling selection. And in this regard, NFV standards could shine at and then become crucial for building a telco cloud system based on multiple functional components, including selected open source ones, yet avoiding coupling with specific implementations to keep the standards open and future-proof for new emerging technologies. Furthermore, studies and analysis in NFV could be enhanced to guide the technology evolution and help network operators to make decisions on when and how to make such changes.

3.3 NFV multi-cloud and multi-technology

The "top-down" network planning will become mainstream in the future thanks to a broader multi-cloud and multi-technology scope of NFV. By building a new distributed network system integrating multiple data centers (at many various and distributed locations), cloud computing, and transport network, service providers will be able to guide the resource requirements of hotspot areas to low-cost areas in an orderly manner. This will optimize the deployment, achieve energy saving, and reduce the overall construction costs of NFV infrastructure thanks to economies of scale. NFV will be enhanced to accommodate various workload requirements and cope with heterogeneous infrastructure avoiding the introduction of siloed systems.

By managing the infrastructure resources of different cloud platforms, such as multiple public, private, and hybrid clouds, and other heterogeneous cloud resources, an NFV-based multi-cloud will shield from the differences between different underlying cloud platforms. This leverages on the fact that NFV infrastructure resources, even if heterogeneous to some extent, can be pooled using NFV technologies and be reused for many kinds of network services deployments, hence becoming a single infrastructure substrate for the network. With a unified NFV-based multi-cloud environment, services will be more flexibly distributed while guaranteeing service continuity between clouds.

Nonetheless, to be able to address telecom networks' high performance and high reliability requirements, especially for data processing and transmission, NFV needs to be further inclusive in adopting the various and multiple types of technologies. Acceleration technologies such as DPDK and SR-IOV, and other related configuration parameters, such as CPU Pinning, non-uniform memory access (NUMA), etc., are widely used and supported by NFV standards. But as the edge computing and utilization of big data analysis and machine learning expand, demand for acceleration technologies and use of high-performance network devices, such as SmartNICs, is ever increasing. The modeling and management of specialized acceleration devices such as GPUs, NPUs and DPUs needs to be considered in developing future NFV standards, to achieve a unified and optimized substrate of acceleration resources to support these new business cases.

In the ubiquitous deployment scenario at the edge, heterogeneous hardware dedicated to edge nodes, can be used as lightweight NFVI edge resources. Vendors of edge infrastructure may use different technologies to meet the requirements of different industries, for example, X86 or ARM based processors used for different network function workloads, GPU or DPU used for video processing and machine learning, SmartNIC or NPU used for heavy traffic processing and forwarding. In addition, in an access device (e.g., DU), the plug-and-play infrastructure resource boards can be built in to provide NFVI heterogeneous hardware resources for end users, meeting the future ultra-low latency service requirements. Furthermore, it will not be possible to neglect the rise of new types of infrastructure beyond typical compute, storage, and network resources, such as programmable metasurfaces, which will not only provide new opportunities for extending the outreach and modeling of network services, but also become an intrinsic part of the infrastructure to be managed and/or exposed to NFV systems.

Looking into the future, OS container virtualization might not be the end of the journey. Alternative solutions like unikernels, in-kernel VMs, serverless, WebAssembly, etc. are also emerging, which could be leveraged to cover existing and new use cases, as well as facilitate further decoupling of software (for network functions and other applications) from hardware. The NFV technologies need to evolve to support the changing needs from telecom service providers and vendors, and ideally, as initially intended, create a truly independent virtualization architecture and solution.

3.4 Unified NFV management and orchestration

With the widespread introduction of more heterogeneous hardware resources (as introduced in the previous section), to avoid a fragmented management of these resources and reduce the difficulty for services to use them, the NFV framework will need to become the unifying "glue". It is envisioned that the NFV framework will provide unified management for heterogeneous hardware and become the "multi-cloud" management platform. By establishing an abstract model of heterogeneous hardware, unifying infrastructure resource orchestration, and providing standard API, the NFV will cope with the differences between various forms of diversified heterogeneous hardware. In turn, this will simplify the application development and deployment, as applications will only need to focus on a service logic expression and will not need to care about the differences among underlying hardware. NFV standards should also consider improving the management of "management data" to enable easier and consistent data processing in a multi-vendor, multi-component network environment.

Similarly, as with the case of hardware resources, heterogeneous virtualized infrastructures are also being introduced into the NFV framework. The NFV framework has evolved to support the management of both VM and container However, with the development of new virtualization technologies, the NFV framework will need to cope with different types of virtual infrastructure and, consequently, need to provide a unified infrastructure management framework for them as well.

The maturity of the PaaS technology is becoming a reality, and in some cases thanks to the availability of related open source solutions. While in many cases those solutions have not been envisioned for particular telecom use cases and be rather focused on their use by IT applications, NFV standards should help bridge the gap between IT-based PaaS and telecom PaaS.

A new generation of telecom PaaS platform based on the principles of application-to-infrastructure decoupling, platform services reusability and sharing, and automation will be one of the future driving NFV pillars. As the telecom PaaS platform is enriched with additional functionality, not only concerning to resource-based services (such as storage, networking, etc.), but also to OAM (such as logging, configuration, etc.), applications development will be greatly simplified. This has the potential, on the one hand to ensure further decoupling of the application from the underlying infrastructure, and on the other hand to speed up the application development and shorten the time-to-market to deliver new telecom services.

Finally, an NFV-based multi-cloud management will also need to integrate various OAM supporting tools. In a digital era, consumers around the world have put forward high expectations for uninterrupted high-quality telecom networks. Towards building a highly reliable (five-nines availability or higher) NFV system, key issues such as disaster tolerance capability, redundancy design, predictive management, fault location, cross-layer operation and pooled resources need to be further studied in the next decade. Meanwhile, telco cloud is also expected to leverage new IT reliability-improving concepts such as site reliability engineering, Infrastructure as Code (IaC), chaos engineering, cloud-native observability, and corresponding tools.

3.5 Autonomous networking, automation, and AI

Since the emerging of NFV, automation has been one of the main design goals. The current NFV standard already provides relatively complete support for the lifecycle management of NSs and VNFs. Service providers can deploy, update, and delete VNFs and NSs using standardized templates and APIs, and even support auto-scaling and auto-healing. However, in order to maximize the value of NFV technology and improve operational efficiency and resource utilization, integrating further automation into the operation and management system needs to be considered.

To achieve the foregoing objective, the NFV system needs to use new technologies to gradually achieve higher automation. For example, the DevOps and CI/CD concepts could be applied to improve the end-to-end automation capability of the system, and extend the automation support of the NFV system from mere lifecycle management to service design, testing, integration, etc. Furthermore, CI/CD and DevOps can help cope with the dynamic change of code and API. And in the future, with the comprehensive improvement of NFV end-to-end automation capabilities, AIOps and machine learning operations (MLOps) technologies could be further introduced to implement automatic decision-making and optimization.

MLOps could fully explore the core data characteristics of different scenarios, establish a continuous evaluation indicator set in a targeted manner, open up the closed loop of data feedback, closely track the degradation of model performance, and automatically trigger the retraining and re-integration of AI models, so that AI models can continuously adapt to changes on demand.

To this end, NFV systems need to provide standardized indicators and metrics definition and monitoring interface protocols across multiple layers (including the information from the underlying hardware layer/physical resources), provide high-precision, fine-semantics and cross-level correlated training datasets and model evaluation metrics, and use intent-driven end-to-end automation capabilities further applying the MLOps principles.

In this context, by considering further expansion of automation and AI in an NFV environment, four main areas will be greatly benefited from. Firstly, intent-based network management can optimize how NFV-MANO understands users' requirements for the network, e.g., SLAs, and translates them into network policies and actions. For example, in the network planning and deployment phase, the digital twin technology can be used to simulate and verify network resource requirements, and to automatically perform service deployment based on network configuration policies. In the system operation phase, the system automatically monitors the service availability, resource performance, and resource availability in accordance with the specified QoS and SLAs, predicts and reports the service and resource status, triggers to start preventive measures and pre-planning, and finally achieves self-planning and self-optimization for closed-loop operation.

Secondly, NFV will evolve to be capable of performing more proactive infrastructure fault management, shifting from the current reactive mode. NFV-MANO will provide fault prediction and prevention capabilities by introducing machine learning and AI technologies and troubleshoot potential system problems in advance to ensure network robustness. By conducting system operations, NFV-MANO can continuously monitor the status of infrastructure resource pools in real time and will automatically analyze root causes by using AI and knowledge maps based on reported events and predicted service instance problems to automatically delimit and locate faults and isolate faulty resources, and finally achieve rapid recovery of infrastructure faults.

Thirdly, AI-driven capacity planning and optimization will enable NFV to perform a more accurate capacity planning and efficient use of NFV infrastructure. In the planning phase, digital twins are introduced to accurately simulate the precise planning of auxiliary resource pools. In the deployment phase, the AI model is introduced to predict the capacity of the resource pool and automatically expand and reduce the capacity of the network resource pool to optimize the deployment of telecom services in the entire resource pool. In the operation phase, resource capacity is automatically evaluated and continuously optimized for efficient use.

Finally, AI technology can be used to assist in realizing the NFV unified management to optimize resources through intelligent service scheduling. With the increase of edge nodes, the AI technology can be used to predict network load changes, and network resources can be scheduled in a timely manner to save energy and reduce consumption in edge clouds and decrease service provider’s operational costs.

4 Conclusion and summary

Since its inception, the ETSI ISG NFV has driven, through its community and the resulting standards, a major transformation of the telecom industry. A proven track of implementable standards is the foundation for further evolving the NFV framework towards the support of future telecom network generations.

As an open community, the ISG welcomes participation of all industry stakeholders, including SMEs, research communities, universities, etc. This creates the ideal substance for considering innovative breakthrough ideas, concepts, and approaches in an ever-changing and evolving telecommunications context. A feature-based development process, as currently in use by the ISG, is at the same time a perfect means for, in a controlled manner, continuously evolving the NFV framework by adopting new visions and trends which are determined to add value and enrich it.

In addition, a long history of collaboration with other organizations, including open source communities, has placed the ISG as one of the pillars in the digital transformation that telecom networks are undergoing. This is of critical importance since NFV is nurtured by many different other technologies not originally developed by the ISG itself following a borrow-in philosophy. Specifically, as more concepts and technology are being developed within the open source communities, collaboration with these is a key. On the one hand, as an input of new technology and even de-facto solutions. On the other hand, as a field for open source to develop and integrate end-to-end solutions based on the NFV framework that the ISG produces. Both interaction flows will become extremely important in the future. ETSI ISG NFV standards will play the critical role in leading to a better use and integration of the open source in the carrier grade environment of telco operators.

In the past 10 years, challenges have been encountered in developing and applying NFV technology. These should be taken, together with the advances and trends in other technology areas, as new opportunities for further improving and expanding NFV as a technology foundation for deploying and managing telecom networks. Towards the next decade, NFV standards and systems will embrace more declarative and intent-based management and offer a truly unified management and orchestration framework on an expanded multi-cloud and multi-technology scope, all together fed by increasing levels of automation through groundbreaking AI and related technologies.

As the industry continuously transforms, new challenging and innovative use cases, requirements and novel ideas will need to be considered. And for this, the ETSI ISG NFV has the means to play a key role in evolving NFV. Get on board this train towards NFV’s next decade.

References

[1] Network Functions Virtualisation - An Introduction, Benefits, Enablers, Challenges & Call for Action, 2012. [Online]. Available at: https://docbox.etsi.org/ISG/NFV/Open/Publications_pdf/White%20Papers/NFV_White_Paper1_2012.pdf

[2] Harmonizing Open Source and Standards: A Case Study of ONAP, 2018. [Online]. Available at: https://www.onap.org/wp-content/uploads/sites/20/2018/03/ONAP_HarmonizingOpenSourceStandards_v2_ac.pdf

[3] Harmonizing Open Source and Open Standards: The Progress of ONAP, 2019. [Online]. Available at: https://www.onap.org/wp-content/uploads/sites/20/2019/04/ONAP_HarmonizingOpenSourceStandards_032719.pdf

[4] Harmonizing Open Source and Standards: A Case for 5G Slicing, 2020. [Online]. Available at: https://www.onap.org/wp-content/uploads/sites/20/2020/03/ONAP_HarmonizingOpenSourceStandards_031520.pdf

[5] ETSI GR NFV 003 (V1.6.1): "Network Functions Virtualisation (NFV); Terminology for Main Concepts in NFV".

[6] ETSI GS NFV 006 (V4.4.1): "Network Functions Virtualisation (NFV) Release 4; Management and Orchestration; Architectural Framework Specification"

[7] ETSI GS NFV-SOL018 (V4.4.1): "Network Functions Virtualisation (NFV) Release 4; Protocols and Data Models; Profiling specification of protocol and data model solutions for OS Container management and orchestration".

[8] Kubernetes® online documents. [Online]. Available at: https://kubernetes.io/docs/concepts/overview/

[9] Harmonizing Open Source and Standards in the Telecom World, 2017. [Online]. Available at: https://www.linuxfoundation.org/blog/blog/new-linux-foundation-white-paper-harmonizing-open-source-and-standards-in-sdn

[10] Network Functions Virtualisation (NFV) - Network Operator Perspectives on Industry Progress, 2016. [Online]. Available at: https://docbox.etsi.org/ISG/NFV/Open/Publications_pdf/White%20Papers/NFV_White_Paper2.pdf

[11] Network Functions Virtualisation (NFV) - Network Operator Perspectives on NFV priorities for 5G, 2017. [Online]. Available at: https://docbox.etsi.org/ISG/NFV/Open/Publications_pdf/White%20Papers/NFV_White_Paper3.pdf

[12] Ray Le Maistre - TelecomTV: Private Networks. [Online]. Available at: https://www.telecomtv.com/content/private-networks/aws-unveils-expanded-private-networks-offering-managed-network-hosting-46745/

ETSI
06921 Sophia Antipolis CEDEX, France
Tel +33 4 92 94 42 00
info@etsi.org
www.etsi.org

This White Paper is issued for information only. It does not constitute an official or agreed position of ETSI, nor of its Members. The views expressed are entirely those of the author(s).

ETSI declines all responsibility for any errors and any loss or damage resulting from use of the contents of this White Paper.

ETSI also declines responsibility for any infringement of any third party's Intellectual Property Rights (IPR), but will be pleased to acknowledge any IPR and correct any infringement of which it is advised.