# Common Platform Architecture for Network Function Virtualization Deployments

Dharma Rajan
VMware, Inc.
Palo Alto, California, USA
drajan@vmware.com

*Abstract*—**Network Function Virtualization (NFV) provides the technology pathway to meet next-generation application needs for mobile cloud adoption. Underpinning this is the virtualization of core infrastructure, such as compute, storage, and networking. Virtualizing key mobile networking functions, such as access networks (4G, LTE, CDMA, non-3GPP Wi-Fi), transport networks (radio access network [RAN], eNodeB, content delivery network [CDN]), and core networks (Evolved Packet Core [EPC]), IP Multimedia Subsystem [IMS], VoLTE), is now reaching a phase where production-level deployments are starting to happen.**

**There are four key pillars of virtualization that form the basis for NFV. The focus of this paper is how NFV deployment can be achieved, using a "Common Platform" approach across access, transport, and core networks with a high-performance hypervisor for compute-, storage-, and network-level virtualization with an integrated operation management system. The software-based approach to providing logical networking functions, such as switching, routing, firewalling, and load balancing, is discussed. From an operations perspective, this paper covers how the virtual infrastructure network will scale, provide high availability, deliver carrier-grade characteristics, and support in managing virtual network function (VNF) with operational support systems (OSS) and business support systems (BSS). Finally, the road ahead, with opportunities that need to be researched, is also presented**

***Keywords— Virtualization, Hypervisor, Logical switching, Logical routing, Distributed firewall, Logical load balancer, VMware vSphere®, Edge gateways, NFV, Cloud, Core networks, CSP.***

## I. INTRODUCTION

Extensive adoption of virtualization by enterprise IT globally is proven. Service providers have also virtualized many workloads in the telecom environment. Gartner [1] specifies that about 75 percent of x86 server workloads are virtualized today. Work done by early adopters like Mobile Cloud Networking [2] and experience gained by enterprises have encouraged CSPs to adopt virtualization technologies in access, transport, and core networks at volume. The European Telecommunications Standards Institute (ETSI) Network Functions Virtualization (NFV) Industry Specification Group (ISG) has published specifications on NFV Infrastructure (NFVI) [3], NFV hypervisor domain [4], and NFV Performance & Portability best practices [5] that form key baselines for NFV adoption by CSPs. One can implement a solution conforming to the framework proposed by ETSI ISG NFV by using the VMware vCloud® NFV™ platform [6].

Telecom and NFV application workloads are different from typical enterprise application workloads. Enterprise workloads are more Platform as a Service-based (PaaS). Telco and NFV workloads are more Infrastructure as a Service-based (IaaS). Generally, many of the telco workloads tend to be any combination of latency-sensitive, jitter-sensitive, or demanding of high packet rate throughputs or aggregate bandwidth, and therefore must be tuned for best performance at the hypervisor level. Thus, it is critical that compute, storage, and network virtualization can collectively guarantee and meet the requirements for CSPs. This paper positions a "Common Platform" approach to build a vCloud NFV environment for use across access, transport, and core networks, along with a virtualized infrastructure management (VIM) system that meets the ETSI reference architecture (as shown in Figure 1), and enables telecom service providers to deploy many types of VNFs. Let us take a detailed look at each of these areas.

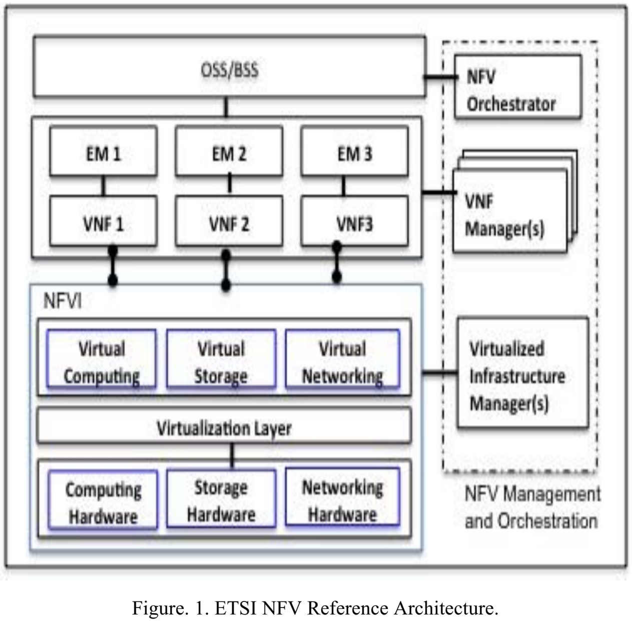

Figure. 1. ETSI NFV Reference Architecture.

978-1-5090-1754-6/16 $31.00 © 2016 IEEE
DOI 10.1109/MobileCloud.2016.10

IEEE computer society

## II. Common platform compute Virtualization

During the last decade, the telecom industry has moved from proprietary hardware to telecom-specific commercially off-the-shelf (COTS) hardware based on the Advanced Telecommunications Computing Architecture (ATCA) [7]. Now there is a shift happening to move to standard x86-based, cost-effective, commodity servers. In the enterprise IT segment, x86 server-based virtualization adoption has been well proven in enterprises of all sizes, with the federal and public sectors driving the "virtualization first" approach in production environments. Today, approximately 50–60 percent of business critical applications are virtualized. More than half a million customers have adopted the VMware vSphere hypervisor (VMware ESXi™) virtualization platform in production. Service providers globally have started to use the vSphere platform to deliver production-grade virtualized mobility service covering millions of subscribers. In the last few years, vendors like Intel and AMD have added support in hardware for virtualization extensions. This has helped improve performance and scale for hypervisors to support more workloads with higher consolidation ratios. Thus, compute virtualization, which is the first pillar, has proven to be ready for adoption by telco's for NFVI use.

ESXi is the common compute virtualization platform for NFVI, as shown in Figure 2. VMware vCenter Server™ is a service that acts as a central administrator for ESXi hosts connected in a network. vCenter Server lets you pool and manage the resources of multiple hosts. The vSphere platform provides the following key elements for NFV environment use in access, transport, and core networks.

a) *Clustering* – A cluster is a collection of ESXi hosts and associated virtual machines (VMs) that have shared resources and are managed by the VMware vCenter™ management system. Clusters can be enabled for VMware vSphere High Availability (vSphere HA), VMware vSphere Fault Tolerance (vSphere FT), and VMware vSphere vMotion®.

b) *CPU Virtualization* – When an ESXi host runs multiple VMs, it allocates to each VM a share of the available physical resources. With default resource allocation settings, all VMs associated with the same host receive an equal share of CPU per virtual CPU. CPU virtualization emphasizes performance and runs directly on the processor whenever possible. The underlying physical resources are used whenever possible, and the virtualization layer runs instructions only as needed to make VMs operate as if they were running directly on a physical machine.

c) *Resource Pooling* – Resource pools are used to partition the CPU and memory resource of ESXi hosts. Resource pools are the preferred way to guarantee resources to VMs. VM resources (such as CPU, memory, and disks) can be specified with shares, limits, and reservations based on VM workloads.

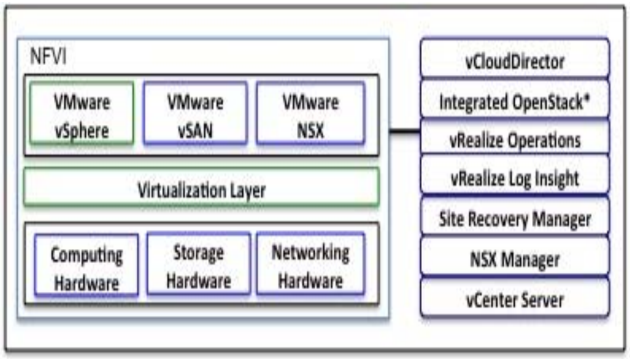

Figure. 2. Common Platform NFV Architecture with Management Stack.

d) *vSphere vMotion and VMware vSphere Storage vMotion* – vSphere vMotion provides the ability to migrate a VM from one host to another. Based on service impact requirements that VNFs have to meet, the level of disruption to service can dictate if a live vSphere vMotion migration or a cold migration is suitable. Storage vMotion helps migrate the virtual disk files of a powered-on VM to a new datastore with no disruption, based on the service impact the VNFs can accommodate.

e) *vSphere High Availability* – The vSphere platform provides higher availability, independent of hardware, operating system, and applications. Hosts in the cluster are monitored, and in the event of a failure, the VMs on a failed host are restarted on alternate hosts, thus providing automatic recovery. It is possible to reduce planned downtime, prevent unplanned downtime, and recover rapidly from outages. Administrators can perform faster and completely transparent maintenance operations without being forced to schedule inconvenient maintenance windows. VMware vSphere App HA helps to detect and recover from application failure through policy-based, application-level monitoring and automated remediation.

f) *vSphere Fault Tolerance for Continuous Availability* – vSphere Fault Tolerance provides a higher level of availability, allowing users to protect any VM from a host failure with no loss of data, transactions, or connections. vSphere FT uses a lock-step technology to provide a replica (secondary) VM running on a different host that becomes active, with its entire state preserved, when the primary VM fails. Thus, vSphere FT provides continuous availability by verifying that the states of the primary and secondary VMs are identical at any point in the instruction execution of the virtual machines. With transparent failover, there is no data loss and network connections are maintained. In addition, affinity and anti-affinity rules are highly important and can be established to allow or prevent VMs from running on predefined hosts.

g) *VMware vSphere Distributed Switch™* – A vSphere Distributed Switch is a single virtual switch that spans across all associated hosts to provide centralized provisioning, administration, and monitoring of virtual networks. This allows VMs to maintain a consistent network configuration as they migrate across multiple hosts. The data plane remains locally on every host and implements packet switching, filtering, and tagging type functionality. The management plane is the control structure on the vCenter server system that administers the networking configuration. Data flows from the VM to VMkernel adapters, down to the physical network. Various NIC teaming and load-balancing policies enable the secure and optimized flow of packets.

h) *VMware vSphere Distributed Resource Scheduler™ (DRS)* – DRS is responsible for the initial placement of VMs and migration using vSphere vMotion. Signal processing workloads, which cover all tasks related to digital processing, such as FFT decoding and encoding in a cellular base station, are expected to be very intensive in CPU processing capacity and highly delay-sensitive. DRS can help correctly manage such scenarios with dynamic resource allocation. VMware vSphere Storage DRS™ provides both VM placement and load balancing based on I/O and/or capacity requirements that would conform to defined VM anti-affinity or VM host affinity rules.

i) *VMware vSphere Distributed Power Management™ (DPM)* – DPM is used to decrease the number of powered-on ESXi hosts in a cluster. DPM works with DRS to monitor resource utilization, and will manage host power on/off by appropriately performing a vSphere vMotion migration of VMs when enough CPU and RAM capacity is available on a cluster, thus saving power and OPEX without operational impact. For latency-sensitive applications, any form of power management adds latency to the path when an idle system responds to an external event. The BIOS P-State and C-State can be used to save power.

j) *VMware vSphere Data Protection™* – vSphere Data Protection provides disk-based backup of virtual machines and applications. It provides agentless VM backup and restore capabilities, automated backup verification, direct-to-host emergency restore, file-level restore, and checkpoint and rollback mechanisms. All these features collectively provide the highest level of confidence in backup data integrity. It also reduces complexity and deployment time, which is very important for CSPs.

## III. Storage Virtualization

The key challenge with storage today is that it is built on expensive proprietary hardware, provisioned in silos for each application, leading to overprovisioning. In addition, storage infrastructure is normally underutilized from a resource and I/O perspective. With software-defined storage embedded in vSphere, the storage resources are created as pools of HDD or all-flash in a shared datastore, and managed through a storage policy-based management framework. This hyper-converged architecture, as shown in Figure 3, provides data persistence delivered from the hypervisor. It provides high performance through flash acceleration and is highly resilient. The new approach provides policy-driven automation and common management across heterogeneous arrays, as well as dynamic control. Thus, it enables intelligent storage placement at scale, dynamic adjustments in real time, automated policy enforcement, very high IOPS, and consistent performance with sub-millisecond latencies. These are all characteristics that make storage virtualization very attractive for NFV use. Entities like Home Subscriber Server (HSS), Home Location Register (HLR) database, and data plane workloads that have high I/O and memory read/write operations can extensively make use of virtual storage infrastructure.

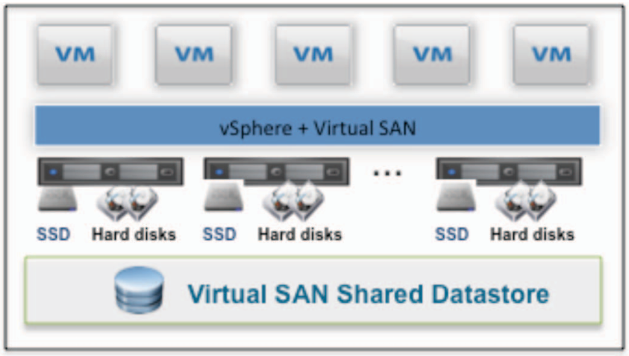

Figure. 3. Virtual SAN – Simple Hypervisor-Converged Storage.

## IV. Network Virtualization

The third pillar of virtualization that helps make NFV a reality is networking and security virtualization. The ability to create virtual networks that reproduce the Layer 2 – Layer 7 network model in software allows for complex multitier network topologies to be created. These virtual networks can be provisioned programmatically in seconds using VMware Network Virtualization Platform (VMware NSX®) [8], making it easy and cost-effective for CSPs to deploy mobile cloud services. VMware NSX simplifies and speeds up the provisioning and management of networking resources through policy-driven automation and hypervisor extension modules, as shown in Figure 4. It helps CSPs move away from legacy/custom hardware and reduces the need to upgrade the entire network hardware stack every few years. VMware NSX helps avoid vendor lock-in scenarios and delivers a new operational model for networking that breaks through current physical network barriers, and enables service providers to achieve better speed and agility with reduced costs. Securing every VNF instance in a multi-vendor NFV platform and securely micro-segmenting with very granular rules between VMs is possible using VMware NSX.

The following list shows four key areas where network virtualization becomes very important for CSPs and drives the best value for mobile cloud and NFV applications.

a) *Logical Switching* – A cloud deployment has a variety of applications across multiple tenants. These tenants require isolation from each other for security, fault isolation, and to avoid IP address overlapping issues. A logical switch creates logical broadcast domains or segments to which an application or tenant VM can be logically wired. This allows for flexibility and speed of deployment while still providing all the characteristics of a physical network's broadcast domains (VLANs) without physical Layer 2 sprawl or spanning tree issues. The physical infrastructure does not have to deal with MAC or forward information base (FIB) table limits because the logical switch contains the broadcast domain in software. A logical switch is mapped to a unique Virtual eXtensible Local Area Network (VXLAN) [9], which encapsulates the VM traffic and carries it over the physical IP network. By using a Layer 2 bridge between a logical switch and a VLAN, one can migrate virtual workloads to physical devices with no impact on IP addresses.

b) *Logical Routing* –Distributed logical routing at the hypervisor level provides dynamic routing between Layer 2 broadcast domains. This allows more direct VM-to-VM communication without needing multiple hops for packets to reach an adjacent VM. Support for dynamic routing protocols (IS-IS, OSPF, BGP), and ECMP, enables efficient East-West traffic routing between VMs in access, transport, and core networks. This improves performance by reducing route cost with reduced hops, increases network efficiency and scale, and enables efficient utilization of the physical router, all of which reduces CAPEX and OPEX.

   The shortest distance to reach a subscriber for voice and video traffic, and offloading traffic with little delay and jitter, is important. Logical switching and distributed logical routing help achieve this for the mobile cloud in a very efficient manner. The user plane workloads (which have higher packet throughputs and carry all user traffic) and media plane workloads (which have moderate packet rates) can be segregated into their own logical switching and routing entities.

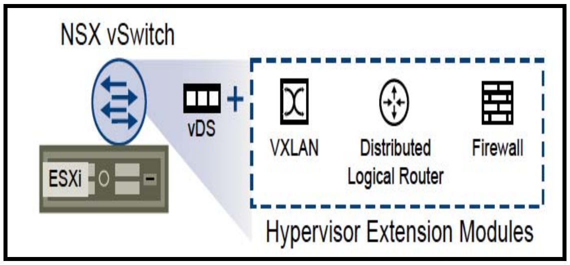

Figure. 4. Logical Networking Modules at Hypervisor Level.

c) *Logical Firewall* – Security is very critical in a virtualized environment, where the surface area of attack is large, if a threat were to enter a service provider data center or CSP's network in any form. The distributed firewall is a hypervisor kernel-embedded firewall that provides East-West traffic protection for VMs, and the NSX Edge™ firewall focuses on North-South traffic (to/from inside to Internet facing side) enforcement at the tenant. The hypervisor-embedded nature of the firewall delivers close to line-rate throughput to enable higher workload consolidation on physical servers. Micro-segmentation, a very granular level of security using isolation based on the zero-trust principle, can be applied at the VM level to tightly secure the environment. The logical edge firewall provides perimeter security functionality including firewall, Network Address Translation (NAT), site-to-site IPsec, and SSL VPN functionality.

d) *Logical Load Balancing* – The load-balancing functionality enables network traffic to follow multiple paths to a specific destination. It distributes incoming service requests evenly among multiple servers in such a way that the load distribution is transparent to users. Load balancing thus helps in achieving optimal resource utilization, maximizing throughput, minimizing response time, and avoiding overload. With support for multiple load-balancing algorithms, acceleration enablement, and integration with third-party load balancers; very efficient load balancing for the traffic in CSP networks can be achieved.

## V. VIRTUAL INFRASTRUCTURE MANAGERS

One of the most important pillars of success for NFV will be the management of the virtual infrastructure and end-to-end management of all applications and services. The Network Functions Virtualization Management and Orchestration (NFV-MANO) architectural framework identifies a Virtualized Infrastructure Manager (VIM), NFV Orchestrator (NFVO), and VNF Manager (VNFM) as management layers.

The VMware vCenter and NSX Manager provides inventory management for the compute, storage, and networking virtualized infrastructure. The resource management helps to manage resource allocation to workloads in the multi-tenant environment. With NFV, secure multi-tenancy at the network level can be established with VMware NSX logical switching and micro-segmentation. VMware vCloud Director®, as an abstraction layer, enables multiple NFVI tenants to be deployed using managed templates for self-service tenant provisioning, and manages tenant resource consumption by throttling the right resources, providing fairness, with network isolation and fine grained security controls. Through orchestration, vCloud Director interfaces can connect to Northbound management and orchestration components, such as VNF Managers.

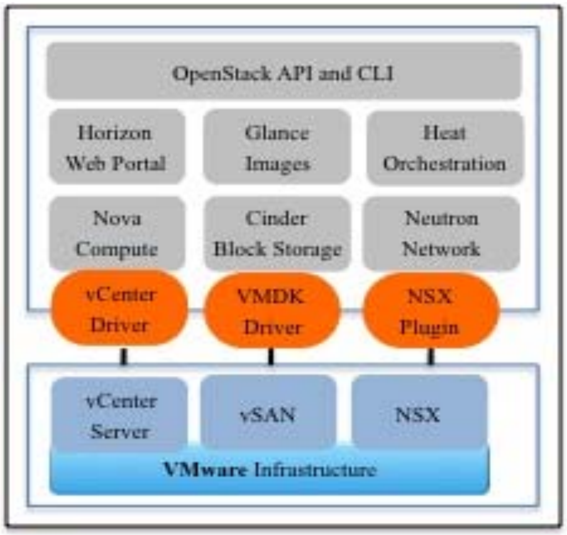

Figure. 5. VMware Integrated OpenStack Framework.

Management of NFVI can also be accomplished using the VMware Integrated OpenStack [10] distribution that integrates with the NFVI base through open source drivers and plugins, as shown in Figure 5. OpenStack provides the abstraction layer to consume the powerful VMware optimized production-grade common platform architecture. Complete monitoring and troubleshooting from OpenStack to the infrastructure layer can be accomplished using management packs built for VMware vRealize® Operations Manager™ and VMware vRealize Log Insight™.

Day 2 operations are very critical. Every network element, element management systems (EMS), OSS, and BSS needs visibility to operational activities that cover all OAM&P areas, such as faults, configuration, accounting, performance, and security. The desire for unified management across traditional and new cloud applications by CSP must be met. For this, intelligent health monitoring, self-learning predictive analytics, smart alerts, and self-healing techniques are used to identify risk situations. Issues are thus remediated before users are impacted by the operations management system. The capability to set dynamic thresholds that adapt to workload changes and eliminate alert storms and false positives exists in vRealize Operations Manager. Super metrics that combine hundreds of KPIs into health, risk, and efficiency scores help with problem detection from multiple symptoms and provide recommendations for proactive actions, making the system highly reliable for use.

The management framework, shown in Figure 6, offers a “single pane of glass” view to monitor and manage the state of the infrastructure. vRealize Log Insight helps collect logs, alarms, and events, and correlates them from different infrastructure layers, performing various types of analytics in real time. This helps with auto-correction and self-healing capabilities that provide faster ways to identify and resolve issues, reducing issue resolution time and improving overall efficiency of day-to-day operations.

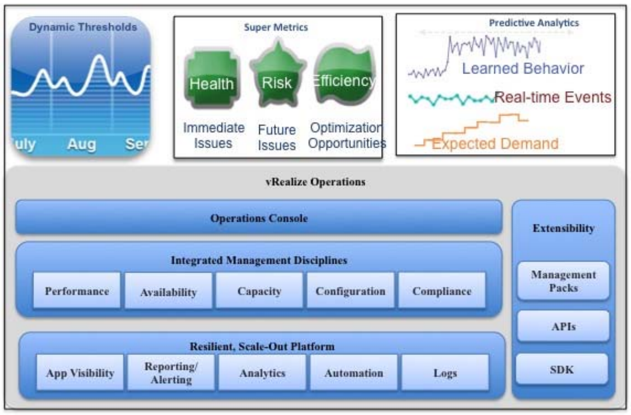

Figure. 6. Operations Management Framework.

The vSphere virtual infrastructure, vCenter management, and other operations, automation, and orchestration tools in the platform provide GUI, CLI, and OSS interface connectivity with RESTful APIs that enable CSPs to easily implement virtualization in the access, transport, and core networks. Seamless operational activity, such as patching, upgrade, troubleshooting, monitoring, diagnostics, SNMP, and syslog support, is available with the platform. This makes it easy to maintain and perform diagnostics and meet carrier-class performance, availability, and scalability requirements for a virtual network that is compliant with NFV MANO documentation.

## VI. DISASTER RECOVERY

Business continuity is very critical in the event of a disaster, which could take the form of a partial outage or a full outage due to planned or unplanned events. The expectation from any mobile customers is a minimum of five-nines availability of network at all times. Providing geo-redundancy is common for telcos. Thus, it is critical that VMs are protected, along with the data, at all times. Traditional disaster recovery plans depend on a very complex set of processes and infrastructure: duplicate data centers, duplicate server infrastructures, processes for getting data to a recovery site, processes for restarting servers, processes for reinstalling operating systems, and so on. Because disaster recovery can be complex, organizations often find themselves unable to provide good protection to more than a privileged few of their production workloads, leaving other workloads unprotected or poorly protected.

To minimize data loss in the event of a disaster, hypervisor-based automated data replication is initiated. To recover quickly from disaster, as shown in Figure 7,

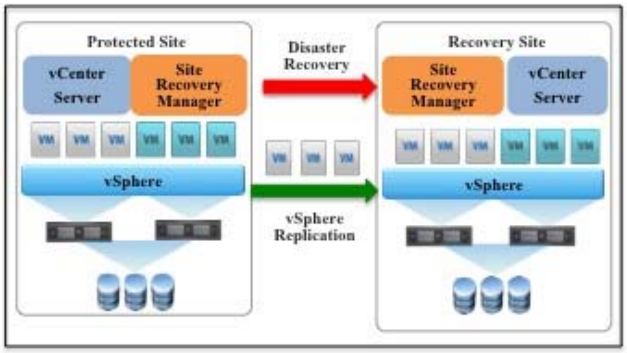

Figure. 7. Automated Disaster Recovery.

VMware Site Recovery Manager™ can be used between primary and recovery sites. With VMware NSX integration into the disaster recovery process, specific networking objects are replicated. Thus, in the event of a disaster, the need for changing IP addresses does not exist, and business continuity is achieved without any network changes. This dramatically simplifies the setup and ongoing management of recovery and migration plans. It enables users to replace traditional, complicated, manual runbooks with centralized recovery plans. This can cut down the time required to set up a recovery plan from weeks to minutes. Automating the failover and migration process allows users to test their recovery plans non-disruptively as frequently as required to measure their recovery time objectives (RTOs) and test their ability to identify any issues.

The failover process is entirely automated to eliminate errors inherent with manual processes. With the ability to automatically fail back applications to their original site, entire sites can be recovered after the disaster. Planning for disaster recovery for NFV solutions is key to achieve continuous operations and meet customer SLAs.

## VII. The Road ahead

NFV workloads are more sensitive to latency and jitter when they carry media streams. Some NFV applications (for example, virtualized Evolved Packet Core) process both latency sensitive and jitter sensitive sub-classes of traffic, and are therefore both packet throughput-intensive and latency-sensitive by nature. The opportunity to tune the infrastructure to meet these varying demands by developing intelligent ways to handle traffic and more efficient algorithms exists.

Virtualizing extremely low latency-sensitive applications [11] can be accomplished. It is critical to confirm that all of the good features and benefits that virtualization offers are leveraged in the process without bypassing the virtualization layer. Using mechanisms like DirectPath I/O and enabling single root I/O (SR-IOV) for high-speed packet processing applications might show interim benefits, but it comes with certain costs and risks. Opportunities exist to find a way to efficiently use the virtual infrastructure to build these applications.

Maintaining feature set compatibility, addressing the stringent performance requirements of the same functions as in non-virtualized networks, and meeting emerging application needs are ongoing investigations.

Integration with legacy environments is going to be a requirement from CSPs that need to be supported until full migration to a virtualized infrastructure happens. The ability to do this without configuration changes is key for success.

Establishing a seamless migration strategy from legacy to cloud with minimal to no downtime, meeting carrier-grade networks needs with a minimum of five nines of availability, and providing a good disaster recovery strategy will enable smoother operation of the environment.

Finally, a big key to success will be operational transformation. This includes assimilating technology and implementing people and process changes needed to operate and manage the virtual infrastructure.

## References

[1] Gartner report - Magic Quadrant for x86 Server Virtualization Infrastructure, July2015, http://www.gartner.com/technology/reprints.do?id=1-2JGMVZX&ct=150715&st=sb

[2] Mobile Cloud Networking, http://www.mobile-cloud-networking.eu/site

[3] ETSI GS NFV-INF 001: NFV Infrastructure Overview, http://www.etsi.org/deliver/etsi_gs/NFV-INF/001_099/001/01.01.01_60/gs_nfv-inf001v010101p.pdf

[4] ETSI GS NFV-INF-004: NFVI; Hypervisor Domain, http://www.etsi.org/deliver/etsi_gs/NFV-INF/001_099/004/01.01.01_60/gs_nfv-inf004v010101p.pdf

[5] ETSI GS NFV-PER 001: NFV Performance & Portability Best Practices. http://www.etsi.org/deliver/etsi_gs/NFV-PER/001_099/001/01.01.01_60/gs_nfv-per001v010101p.pdf

[6] vCloud Network Function Virtualization (NFV) http://www.vmware.com/industry/telco/overview.html

[7] PCI Industrial Computer Manufacturers Group (PICMG) AdvancedTCA Overview
https://www.picmg.org/openstandards/advancedtca

[8] VMware Network Virtualization Platform http://www.vmware.com/products/nsx/

[9] Virtual eXtensible Local Area Network (VXLAN) https://tools.ietf.org/html/rfc7348

[10] VMware Integrtaed Open Stack http://www.vmware.com/products/openstack/

[11] Deploying Extremely Latency-Sensitive Applications in VMware vSphere 5.5 http://www.vmware.com/files/pdf/techpaper/latency-sensitive-perf-vsphere55.pdf