Reducing The Tradeoff Between Performance And Management Using Container And Cloud-Native Approaches

A Technical Paper prepared for SCTE•ISBE by

Karthik Krishna
Senior Solutions Manager
VMware
Palo Alto, California
kakrishna@vmware.com

Jambi Ganbar
Director, Telco Partner Solutions
VMware
Berlin, Germany
jganbar@vmware.com

Table of Contents

Title **Page Number**

List of Figures

Title **Page Number**

Introduction

Virtualization delivers efficiencies and creates new capabilities, while expanding the network cloud and blurring boundaries at the network edge. In the initial wave of virtualized functions and appliances, demonstrating that software-based infrastructure could be as robust as its hardware equivalent has been a paramount concern. Recent developments in data plane acceleration technologies, including SmartNICs, are playing an important role in meeting and exceeding performance requirements. Going forward, as services expand to meet opportunities, especially in the evolving network edge, the challenge will be to narrow the gap between performance and manageability. A Next-Gen Cloud model promises to minimize those trade-offs.

1. Business and Technology Drivers

1.1. Demand for Services

For change or evolution in a network to occur, there needs to be a business reason. A prerequisite for any change is demand for services. There also needs to be technology that can make the delivery of those services more valuable and profitable. But in regard to revenue generation, the cable industry is in a relatively strong position.

Demand for high-speed data remains strong. Since March 2020, subscribers have especially taken advantage of the upstream capabilities of DOCSIS networks. The industry's flexible content delivery networks (CDNs) are an attractive platform for businesses seeking to deliver automated and targeted advertising. Cable's telephony business has declined, as have all landline offerings, but the industry's MVNO business is up. There is also a growing opportunity to work with mobile operators on backhaul, mid-haul and fronthaul (Xhaul), as smaller cells proliferate in the 5G/LTE-A transition. MSOs continue to build their business services portfolio, both in the SMB market and beyond.

1.2. Edge Computing/IoT Opportunities

Among those existing services, perhaps the work being done in CDNs (manifest manipulation, formatting, encryption, etc.) is the leading example of computing at the network edge today. But at the intersection of wireless, consumer and B2B, new applications involving mobile and IoT devices that need similar services are emerging across numerous verticals. A few examples include:

- Cloud gaming – A billion-dollar industry that has grown during the pandemic, cloud gaming benefits from low latency associated with edge-computing; in effect, it "needs an edge."[1]
- Surveillance, mapping – Distributed network sensors from security cameras, drones and IoT devices drive massive volumes of traffic across constrained wireless networks; one opportunity is for more integrated and AI-driven management.[2]
- Health care – From medical-grade wearables to RFID tags in hospital inventories, IoT is playing a key role in today's increasing digital healthcare system; it needs secure, robust connectivity.
- Smart cities – Municipalities aiming to deploy outdoor Industrial (I)IoT systems that connect remote sensors to network servers could leverage vast HFC networks.[3] From automated traffic lights to self-driving vehicles, what's needed is very low latency and very high reliability.
- Augmented reality – Related to gaming, AR applications are also playing a role in business and industrial settings; like other examples, it could benefit from low-latency and AI.

Cable executives see potential in several of these and other categories. But according to a survey conducted by Broadband Success Partners, they face obstacles, including a lack of operational support for monetization and the existing network itself. When asked what is most needed to make edge computing happen, they have one leading answer: "Network infrastructure (physical and virtual) that is programmable and provides network performance information."[4]

1.3. Technology Evolves, Virtualizes

The industry's technology leaders have adeptly exploited the capabilities of the HFC plant with iterative versions of DOCSIS, while anticipating areas of growth, such as the need to meet demands on the upstream or low-latency requirements for Xhaul and other emerging services.[5][6]

Some MSOs have also begun to adopt cloud computing and virtualization, seizing opportunities for efficiency, innovation, lower OpEx, automation, improved management, elasticity and scale. The use cases involve program guides/UIs, CPE devices, CDN and most notably the CMTS/CCAP. In new cases following the distributed access architecture (DAA), the virtualized (v)CMTS is not only disaggregated, but its core software is loaded onto x86 servers and compute power deployed on remote PHY devices (RPDs) in the last mile.

The DAA initiative was driven less by new market opportunities than the need to manage costs, especially those involving power and headend/hub real estate. The goal has been to grow without launching a massive rebuild or upgrade cycle; in other words, to keep cost per bits stable while meeting escalating demand and maintaining stringent service performance requirements.

At the same time, the vCMTS has opportunities at the network edge. Here, again, the industry is being proactive.[7] One relevant trend is the move away from scale-out homogenous servers using x86 CPUs only, to scale-out heterogenous services that include x86 CPUs, as well as graphics processing units (GPUs), field-programmable gate arrays (FPGAs), and other elements.[8]

As usual, prospects arrive with additional challenges. Delivering services over virtualized and software-driven infrastructure within tight performance and cost parameters is no small feat. But there are other requirements to consider, such as agility, network scalability, security, operational support, product lifecycle and overall management capability.

2. The Virtualization Story

2.1. IT Legacy

Network virtualization is far from a new technology. The idea of a virtual machine (VM) goes back at least to the mid-1970s.[9] Two decades later, in the tech boom of the late 1990s, the time was right to focus on software that acted like a real computer with an OS, separate from underlying hardware. The first product of VMware, founded in that era, was a hosted hypervisor, which enabled users to set up and maintain VMs on a single physical machine.

This market primarily focused on IT, on business-critical applications such as data base systems and Web applications. The emergence of the Network Functions Virtualization (NFV) initiative in 2012 shifted attention to areas such as routing, firewalls and load balancing, all workloads requiring much higher packet rates and much lower packet loss rates. In the case of VMware, these requirements led to a "re-architecture" of the entire networking stack, from the virtualized network interface controller (vNIC) emulation to virtual switching in the device driver.[10] As it happens, there were ongoing industry initiatives collaboration to support such efforts.

2.2. DPDK and Fast Packet Processing

The Data Plane Development Kit (DPDK), founded in 2010 by Intel, became an important resource. Made available under an open source license, DPDK was created to be a "vendor-neutral software platform for enabling fast packet processing, upon which users can build and run data plane applications."[11] The open source community at DPDK.org launched in 2013, and the initiative went under Linux Foundation management in 2017. Alternatives to DPDK are available; according to software development community StackShare, they include Beats, Riemann, LibreNMS, PRTG and Nagios XI .

For its part, the DPDK and its resources, which include libraries and drivers to implement network functions in commodity servers with commodity NICs, have driven industry-wide data-plane efficiencies. Partly inspired by DPDK techniques, VMware built and released an initial enhanced networking stack (ENS) that delivered a packet forwarding rate, in 64-byte packets, with 4 times greater efficiency, while maintaining a packet loss rate of less than 0.001 percent. Since then, performance has continued to improve.

2.3. Architectural Considerations

Other developments impacted the landscape. At the OS-level, virtualization enabled the delivery of resource-efficient software in packages called containers, which are associated with a release by Docker in 2013. Soon thereafter came an open-source container orchestration system known as Kubernetes, designed by Google and now maintained by the Cloud Native Computing Foundation (CNCF). While often pitted against each other, these approaches can all coexist. The latest VMware platform, for instance, allows using any combination of VMs and containers.

At the system level, disaggregated functions mean that control, data and management planes can be deployed across a distributed topology. Where to place what, however, follows basic rules. Because of inherent low latency and advances in processing, edge clouds offer performance advantages for data plane-intensive workloads; while control and management plane components (to the extent that they figure within a virtualized infrastructure) can be centralized with regional and global scope.

One example is distributing compute and virtualization via Flexible MAC Architecture (FMA), a CableLabs initiative. FMA provides the flexibility to place the MAC or compute portion of the CCAP anywhere from the DataCenter to Headend/Hub or in the outside cable plant. The FMA enables placement of the MAC of such a data-plane intensive workload closer to the Edge to address low-latency applications like AR/VR, gaming, or autonomous vehicles and centralize other functions like control and management plane. Such a flexible edge computing model enables smooth integration with wireless.

3. MSO Options

3.1. Legacy Appliance-based

The cable industry has several options regarding virtualization. Whether involving routers, CDNs, firewalls, CMTS/CCAP or other equipment, MSOs can expand their infrastructure as needed via standard appliance-based solutions. This largely status-quo approach, even with disaggregation, poses continued challenges to scaling and incurs operational costs and complexities.

3.2. Virtualization – Evolved on Bare Metal

Virtualizing on bare metal, i.e. on a single-tenant physical server, can deliver certain benefits, including performance and resource utilization; although challenges remain across a number of areas, including scalability, data persistence, networking, security, and management/operations.

3.3. Native on Hypervisor

This second road to virtualization enables a unified, converged architecture, with the benefits of higher availability, densification, scale, multi-tenancy and greater manageability. This "Next-Gen Cloud" approach also enables greater agility. Both virtualization approaches entail breaking with legacy models and acquiring new areas of expertise.

4. Next Gen Cloud

To expand upon the third option above, the Next-Gen Cloud, let's consider its platform architecture, benefits and values.

4.1. Next-Gen Architecture

The Next-Gen Cloud schematically rests upon a standard Network Functions Virtualization Infrastructure (NFVI) and Cloud-Native Network Functions (CNF) infrastructure. (See Figure 1.) Above that is the Virtualized Infrastructure Management (VIM) layer, which enables handling containers and related infrastructure "as a service" (IaaS, CaaS). The Virtualization layer provides resource abstraction for computing, networking, and storage and gives the same experience across VMs and containers.

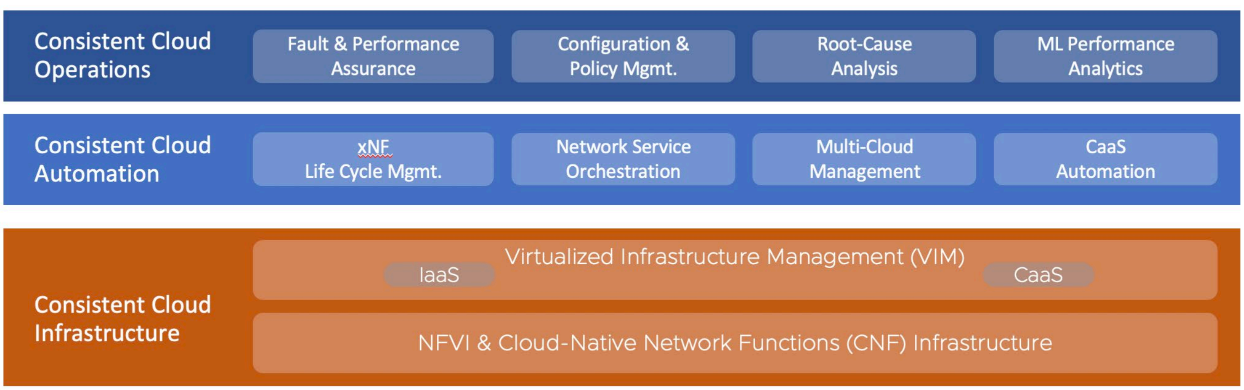

Figure 1 – Next-Gen Cloud Architecture

Cloud automation and operations reside above the infrastructure. Cloud Automation provides multi-layer automation for consistent operations across any network from the core to the edge to the access. Cloud automation enables smooth management and orchestration of the infrastructure layer and automates network functions' placement and lifecycle management over VM- and container-based infrastructure. Cloud Operations provides real-time assurance designed to simplify your network operations through holistic monitoring and performance management, providing comprehensive visibility and automation. Cloud operations in a virtual environment can integrate with existing NetOps tools and data, but they also enable cloud-native ways of fulfilling functions in the areas of fault and performance assurance, configuration and policy management, root-cause analysis (RCA), and machine-learning (ML) analytics.

None of these technologies work without revenue generation and/or positive ROI. Conceptually, at the top of the stack lie a wide range of monetization opportunities; and embedded within the stack are tremendous cost savings and management efficiencies.

4.2. Next-Gen Benefits

The benefits of this approach derive from the interplay of the platform's capabilities and performance in these areas:

Scalability – Orchestration enables elastic scaling of cable workloads, while consistent dimensioning and improved hardware utilization reduces space and power required in headend and hubs. The platform brings uniformity while maintaining security, networking and monitoring requirements.

Operational Efficiency – Having a common platform enables deployment and management of multiple Kubernetes clusters, along with VMs. Sophisticated software, which leads to fully automated services, compares favorably to an inefficient and complicated manual process of managing physical storage devices and using SNMP for monitoring to the NOC. It also provides a consistent experience with a simplified life-cycle management.

Faster Time to Market – This kind of platform allows MSOs to respond quickly to emerging edge compute and other opportunities through an extensive, multi-vendor partner ecosystem, as well as integrated service blueprinting, deployment, monitoring and management.

Workforce Efficiency – Remote deployment and high-performance, self-healing capabilities can reduce truck rolls and the need for on-site service technicians, whether in the access plant, hubs, headends or data centers.

Reduced TCO – Lower CapEx results from increased throughput and performance; reduced OpEx, from software-driven infrastructure and service assurance.

Safe and Easy Convergence – Consistent infrastructure, automation and operations enable easy convergence, with reduced risk of misconfiguration; hybrid VM and container-based applications feature advanced networking and security features.

4.3. Next-Gen Capabilities

Agility. With a hypervisor, a new bare-metal server can connect to the container domain in minutes. That agility is one of the reasons that major public cloud providers use hypervisors to run their container services. Such agility and scalability work because the hypervisor de-couples and abstracts software from the hardware. The Next-Gen Cloud delivers over-subscription capability, improving utilization of the underlying hardware. This compares favorably to the standard practice of racking and stacking, re-segmenting the network and then testing that the segmentation was correct.

Networking. The Next-Gen Cloud implements a single underlay network on VMs to provide end-to-end connectivity and management for both containers and traditional applications. A single underlay network makes it easier to connect containerized applications to traditional, non-containerized components like databases; simplifies network management with centralized policies and advanced security; and enables selecting the overlay network and the service mesh that works best for containerized applications.

Security. Containers and hardware virtualization not only can, but very frequently do coexist and actually enhance each other's capabilities. VMs provide strong isolation, OS automation and an ecosystem of solutions. They enable secure and efficient running of containerized applications in production. Containers being run on hypervisors can take advantage of security innovations, including micro-segmentation, which enables security architects to apply security controls and deliver services at the individual workload level.

Manageability. A comprehensive, flexible platform allows you to deploy and manage multiple Kubernetes clusters as well as to manage, patch and upgrade the container host OS. All these capabilities empower you to run traditional and containerized workloads on a common infrastructure, while ensuring optimal performance and preventing interference between workloads.

Performance. Empowered by vSwitch and Workload Acceleration functions, the Next-Gen cloud's hypervisor can provide more efficient overall workload performance for containers than Linux systems running on physical hardware. The platform uses advanced scheduling algorithms that enable employing modern DPDK packet processing, allocating CPU cycles for efficient networking, and faster packet processing to optimize all workloads. The Next-Gen cloud also employs hardware offloading techniques to dedicate all the compute for workloads and can leverage SmartNIC for infrastructure computing.

5. Data Plane Acceleration

5.1. Performance and Tradeoffs

To look more closely into performance, let's first set the stage. Workloads in service provider solutions involve the control and signaling plane, the data plane, and the management plane. A network element (NE) traditionally was built to handle control and data plane functions in one box, with its own element management system (EMS). With the arrival of SDN, the data and control planes began to separate. Then with NFV, multiple functions within each NE were broken into multiple VNFs, providing service providers with choice, enabling best-of-breed solutions with more software-based control.

Of the three workloads, those on the data plane need to meet the most stringent requirements. (When the data plane is lost, the subscriber notices.) These demands led to the development of software accelerators that put applications needing fast packet processing in the 'fastpath' of a compute host. Management and low-priority applications could go via 'slowpath.' Yet the arrival of virtualization made it difficult to attain expected performance levels, because the hardware used was based upon commercial off-the-shelf (COTS) servers.

To be sure, virtualization correctly done continued to have significant benefits, primarily: those derived from the secure pooling of networking resources across multiple hardware units. A virtualization layer with a virtual distributed switch that spans multiple hosts (servers) can live-migrate workloads and efficiently distribute resources on demand. To gain performance, service providers used two techniques: single root I/O virtualization (SR-IOV), a concept introduced by Intel that allows for bypassing the virtualization layer when it relates to the data path; and DirectPath I/O, which allows for direct access to a physical NIC.

These techniques enabled faster packet processing but came at the expense of not being able to use a wide range of critical features, all provided by the best-in-class virtual infrastructure. Applications built for these pass-through mechanisms also faced security and cloud-ready challenges. Another technique that gained popularity was CPU pinning, which led to locked-in hardware, driving up CapEx and OpEx costs.

Among the tools enabled by the DPDK initiative was a software library that helped enhance performance by allowing for optimized packet allocation across DRAM channels. The principle being that allocation of memory from local nodes and cache-alignment of objects can lead to superior performance. The upshot of these developments was an accelerated vSwitch, now a key to the Next-Gen Cloud model, which overcomes the drawbacks of SR-IO and DirectPath I/O and obviates the need for other techniques, such as CPU pinning.

5.2. Two Data-Plane Acceleration Options

With that background, we can now consider two data-plane acceleration approaches, both involving PCIe NIC devices. The first involves offloading with a standard Performance NIC; and the second, doing so with SmartNIC. Both cases use an x86 host server and the accelerated vSwitch, which supports two configuration modes: Standard, for use with any management or control-plane application; and Enhanced, for use with data-plane intensive applications. In the first approach, the vSwitch is running on the hypervisor, and in the second, it has been offloaded onto the SmartNIC and can leverage separate embedded NIC cores. (See Figure 2 and Figure 3.)

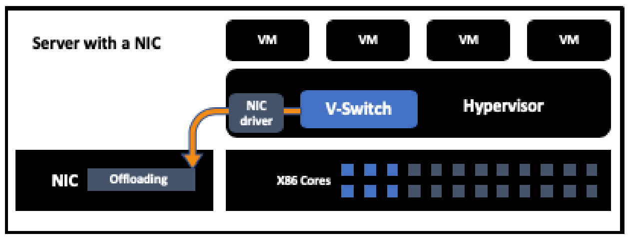

Figure 2 – Performance NIC Architecture

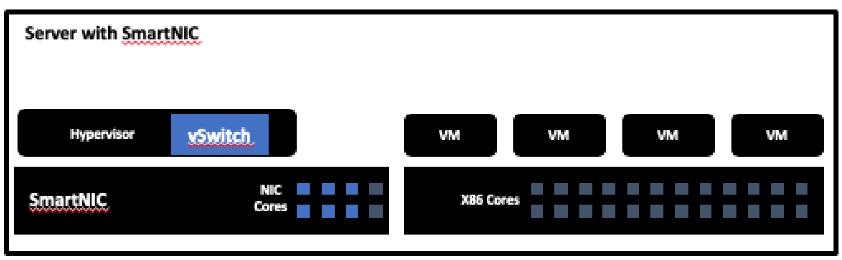

Figure 3 – SmartNIC Architecture

5.2.1. Acceleration with Performance NIC

Operators have several ways to boost performance, not only in the software realm. We will point out five areas. (See Figure 4.) Let's begin with decisions involving hardware elements:

(1) NIC choice – To meet 10G requirements, first choose Performance NICs with TCP Segmentation Offload (TSO) and Checksum Offload (CSUM) capabilities, as well as overlay support.

(2) CPU Choice – Devote a high number of cores for workloads and find the right balance between CPU speed, core count and wattage requirements.

(3) Server architecture – Increased performance is a function of the NUMA balance, choice of the right PCIe and server architecture. Another key consideration is server immutability, in which servers are never modified after being deployed.

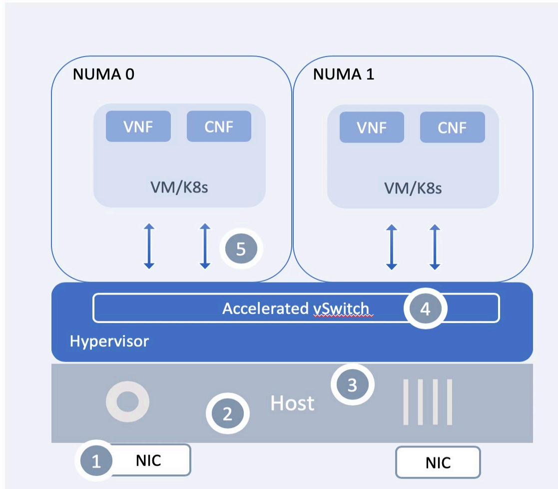

Figure 4 – Acceleration with Performance NIC

With the right mix of hardware, performance can be further enhanced through strategic choices in software implementation (enabled by Next-Gen cloud):

(4) vSwitch Acceleration – Based on Intel DPDK principles, an accelerated vSwitch provides very high throughput, low latency, low jitter, near zero-packet loss, characteristics that are essential to run data plane-intensive workloads, whether at the Core, the Edge or the Access. The accelerated vSwitch achieves higher packet performance by freeing the network to work on any server or modern NIC; dedicating CPU cycles for networking; and deploying faster switching with flow cache, lockless datapath and faster packet processing (SSE). The vSwitch forwards traffic between components running on the transport node (between VMs/VNFs or between containers) or between internal components and the physical network.

(5) Workload acceleration – Other options include using an accelerated vNIC, dedicating (pinned and isolated) vCPUs; topology awareness, and Huge Pages support.

Powered by the accelerated vSwitch and an accelerated workload with the right combination of NIC, CPU and server hardware, this approach can handle the various workloads in use today. VNFs in all types of form factors – whether VMs, containers, or micro services – can run on this kind of enhanced solution. Being agnostic to the application, the vSwitch provides operators with the 'freedom of choice' to pick any VNF that would meet the solution needs.

5.2.2. Performance NIC Test Results

Tests of Performance NICs, which can flag latency- and jitter-sensitive applications and choose selective vCPU pinning, generated results well within acceptable parameters. In a traffic stream starting and terminating with Spirent TestCenter software, with latency that included both transmission delay introduced in forward and reverse paths as well as software processing times, the average latency contributed by virtualization functions amounted to less than 30 microseconds; for jitter, less than 10 microseconds. Both are negligible amounts. When testing throughput of varying packet sizes, the actual throughput approximated the line rate, with a representative bundle of video-heavy packets generating a speed of 3.98 Mpps and 4.5 Mpps, respectively, for 3- and 4-core implementations.

5.2.3. SmartNIC

There is a growing consensus that CPU cores are such a precious commodity that they should never do network, storage, or hypervisor housekeeping work, but rather focus on core computation. That means network offloads and storage offloads need to be mainstream in the coming years, creating an even more asymmetric and heterogeneous processing environment than many are envisioning down the road. Throwing more and more compute at the problem is unsustainable. The goal is to iterate fast, but I/O processing is unable to keep pace with compute processing. SmartNICs enable offloading all or most of a virtual switching stack or a large chunk of a distributed storage stack.

A SmartNIC is a high-performance NIC, equipped with general-purpose compute cores capable of running an OS and general-purpose applications and workloads. It is both a NIC, with flow-match hardware capabilities, and a mini server. In architectural terms, its purpose is to move server management to SmartNIC cores and offload the entire vSwitch/hypervisor. As a result, the host x86 remains dedicated to workloads, increasing its efficiency. The host x86 can run either a VM/container or a bare-metal.

One of the big reasons to use a SmartNIC is not only to save CPU resources, but to scale in different ways, for instance by saving power and space. Offloading the heavy compute enables an operator to free up the CPU for others things, such as multi-access edge computing (MEC) in 5G. Another key to a full CPU offload is the potential to implement policy and priorities straight on the input, without having to use the OS at all, which makes for a powerful story.

Other benefits include a clear-cut between infrastructure and application, air-gap security, cost management, application acceleration (e.g. security, load balancers, etc.). In organizational terms, the SmartNIC clearly splits ownership and enables multi-tenancy. (See Figure 5.) SmartNICs are positioned to handle workloads that are scaling to meet 10G requirements.

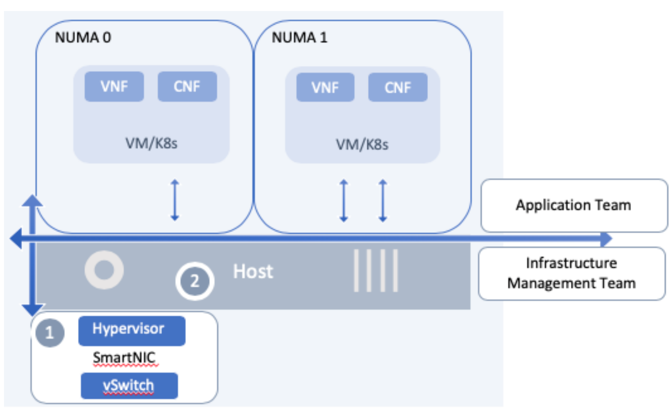

Figure 5 – Acceleration with SmartNIC

Because SmartNICs come equipped with onboard local persistent storage, a large amount of DDR RAM, multi-level caches, PCIe root complex, virtualized device functions and I/O capabilities, they have significant potential for use cases that extend beyond network acceleration and vSwitch offloads. Those cases could include but are not limited to management plane offload and virtual device acceleration.

6. Next-Gen Cloud Reference Architecture

The reference architecture for Next-Gen Cloud includes five tiers: from physical infrastructure to VM and container-based platforms to resource orchestration to cloud automation to solutions. (See Figure 6.) On the solutions tier, the vCMTS is only one of many, which also include L2/L3 solutions, SD-WAN and various video-centric services. Being extensible to emerging solutions and business models is one of the most compelling attributes of the Next-Gen Cloud.

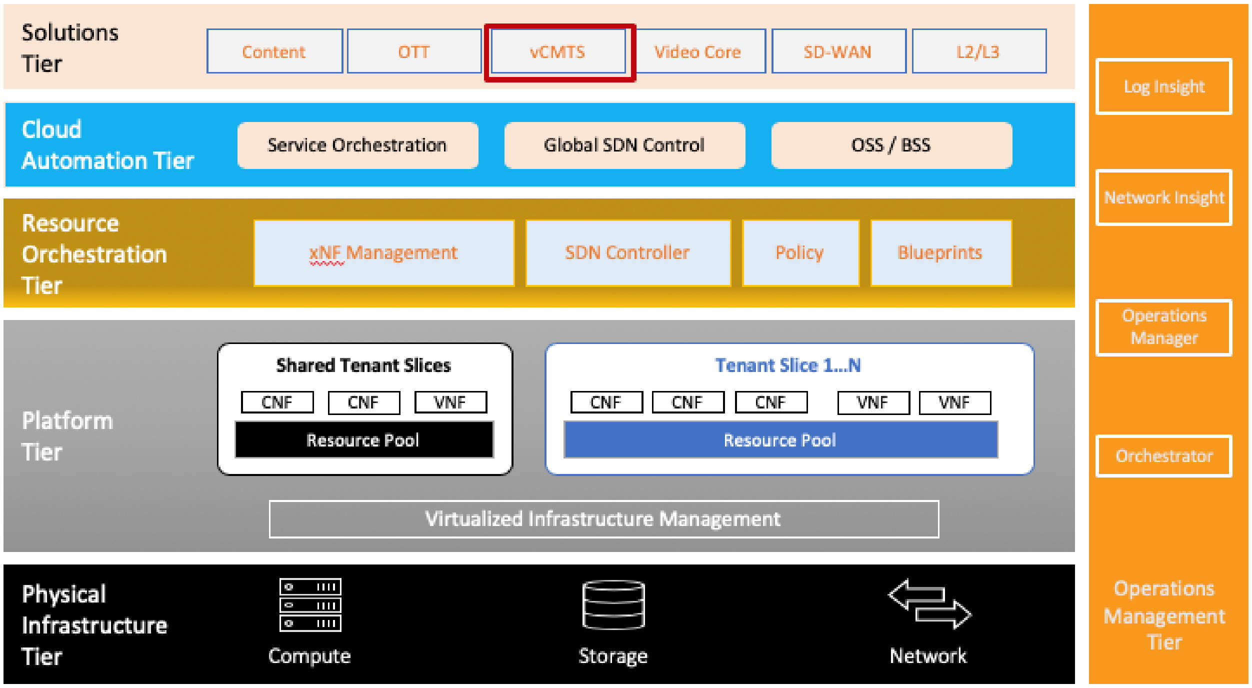

Figure 6 – Next-Gen Cloud Reference Architecture

Moving from that top tier down the stack, we note that it features centralized control and management, including embedded automation and optimization. Its highly flexible infrastructure as code (IAC) model is characterized by a heterogenous runtime with Network Function and Resource Isolation. Finally, it rests upon commodity hardware and storage and leverages vertically integrated

7. vCMTS Use Case

Compared to the mobile industry, where NFV has proved to be effective in 4G LTE production deployments and has become the basis for building 5G networks across extensive footprints, virtualization in the cable industry has taken a more gradual and piecemeal approach. The initial deployments of the vCMTS occurred after MSOs assessed different approaches to DAA and enhancements to DOCSIS. It is beyond our scope to assess those ongoing rollouts, but it is noteworthy that one summary practical lessons from a DAA deployment with a vCMTS last expressed concerns at the operational and management layer, in particular the need for tools.[12]

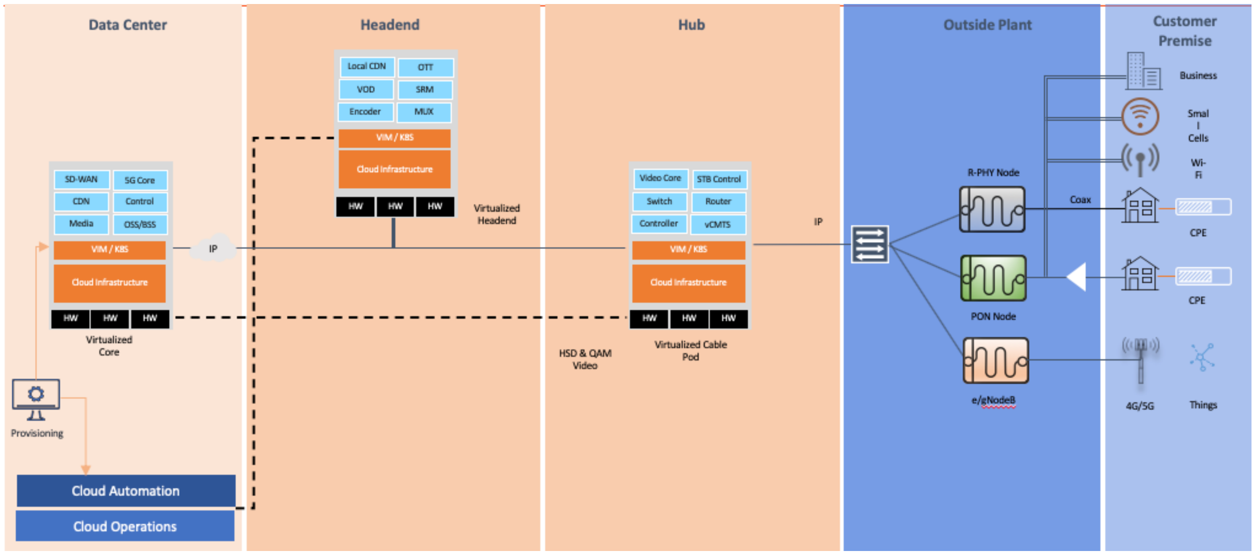

Figure 7 – Scale and Convergence in Next-Gen Cable Cloud

If manageability is an issue with the current vCMTS model, in which the R-PHY node is the only element in the outside plant, will it not become more so if and when MSOs deploy virtualized PON nodes and eNodeB devices? (See Figure 7.) Moreover, in addition to services delivered to residential CPE endpoints, those involving enterprises, small cells, Wi-Fi, 4G/5G and the proliferating number of edge compute business cases also need to be managed. Conducting efficient root-cause analysis over a network with no centralized control and integrated operations management is just one challenge we could mention.

On the other hand, with Next-Gen Cloud model, there is ongoing innovation in management, automation and OSS. Virtualization changes the way we can set up, handle faults, share resources, and more.

Conclusion

Growing demand for connectivity, compute and storage at the network edge is an exciting prospect for MSOs shifting to a distributed and virtualized network infrastructure model. Advanced data plane acceleration that can drive high performance over software-based infrastructure is also an encouraging development. Less exciting is the prospect of building more operational siloes to handle a growing menu of use cases and services, which defeats one of the fundamental reasons for the drive toward virtualization in the first place: "the principle that it is no longer necessary to solve for the platform and runtime layer below network applications in a different and particular way for each additional application – with all of the attracted cost, complexity, and operational management overhead that differentiation implies."[13]

Abbreviations

4G LTE	fourth generation, long-term evolution
5G	fifth generation technology standard for cellular networks
AI	artificial intelligence
AR	augmented reality
CaaS	container-as-a-service

CCAP	converged cable access platform
CMTS	cable modem termination system
CNCF	Cloud Native Computing Foundation
COTS	commercial off-the-shelf
CPE	customer premises equipment
CPU	central processing unit
CSUM	checksum offload
DAA	distributed access architecture
DDR	double data rate
DOCSIS	data over cable service interface specification
DPDK	data plane development kit
DRAM	dynamic random access memory
eNodeB	E-UTRAN node B, or evolved node B
EMS	element management system
ENS	enhanced networking stack
GPU	graphics processing unit
HFC	hybrid/fiber coax
I/O	input/output
IAC	infrastructure as code
IIoT	industrial internet of thing
IoT	internet of things
LTE/A	long-term evolution/advanced
ML	machine learning
MEC	multi-access edge computing
MSO	multiple system operator
MVNO	mobile virtual network operator
NE	network element
NFV	network functions virtualization
NFVI	network functions virtualization infrastructure
NIC	network interface controller (or card)
NUMA	non-uniform memory access
RAM	random access memory
OS	operating system
PCIe	peripheral component interconnect express
RCA	root-cause analysis
RFID	radio frequency ID
RPD	remote PHY device
SD-WAN	software-defined wide area network
SP	service provider
SR-IOV	single root I/O virtualization
SSE	streaming SIMD (single instruction multiple data) extensions
TCO	total cost of ownership
TSO	TCP segmentation offload
UI	user interface
vCMTS	virtual CMTS
VIM	virtualized infrastructure manager
VM	virtual machine
vNIC	virtual NIC

Bibliography & References

[1] Alan Evans, "Why Gaming Needs an Edge," SCTE-ISBE, 2019.

[2] Sandeep P. Chinchali, et al., "Neural Networks Meet Physical Networks: Distributed Inference Between Edge Devices and the Cloud," HotNets-XVII, November 15-16, Association for Computing Machinery, 2018.

[3] Charles Chapman, "How HFC-based Industrial IoT Gateways Improve Performance of Remote IoT Sensors," SCTE-ISBE, Cable-Tec Expo, 2019.

[4] David Strauss, "Cable and edge computing: The rest of the story," Light Reading, May 18, 2020.

[5] Jennifer Andreoli-Fang, John T. Chapman, et al., "Blueprint for Mobile Xhaul over DOCSIS," SCTE-ISBE, 2019

[6] Greg White, et al., "Low Latency DOCSIS: Overview and Performance Characteristics," SCTE-ISBE, 2019

[7] Omkar Dharmadhikari, "Moving Beyond Cloud Computing to Edge Computing," CableLabs, May 1, 2019

[8] Randy Levensalor, "Give Your Edge and Adrenaline Boost: Using Kubernetes to Orchestrate FPGAs and GPU," CableLabs, Jan 28, 2020

[9] Gerald J. Popek, Robert P. Goldberg, "Formal Requirements for virtualizable third generation architectures," Communications of the ACM, July 1974

[10] Moiz Alam, et al., "High Packet Rate Networking on vSPhere," VMware

[11] DPDK Project Charter, dpdk.org/charter

[12] Asaf Matatyaou, "Practical Lessons of a DAA Deployment with a Virtualized CMTS," SCTE-ISBE, 2019

[13] Andrew Bender, "A Roadmap for Virtualization in HFC Networks," SCTE-ISBE 2019

