The What, Why and How of Network Virtualization

Part 1: Introduction and Network Virtualization

By Dr. Jim Metzler, Ashton Metzler & Associates
Distinguished Research Fellow and Co-Founder
Webtorials Analyst Division

Sponsored in part by:

Produced by:

Table of Contents

Executive Summary

Over the last year, the hottest topics in networking have been Network Virtualization (NV) and software defined networking (SDN). There is, however, considerable confusion amongst enterprise IT organizations relative to these topics. There are many sources of that confusion, including the sheer number of vendors who have solutions that solve different problems using different solution architectures and technologies, all of whom claim to be offering SDN and/or NV solutions.

The primary goal of the ***2013 Guide to Software Defined Networking & Network Virtualization*** (The Guide) is to eliminate that confusion and accelerate the adoption of NV and/or SDN. The guide will achieve that goal by walking the readers through the following set of topics:

1. What are the problems and opportunities that NV and SDN help to address?
2. What are the primary characteristics of NV and SDN solutions?
3. How does NV and SDN help IT organizations respond to problems and opportunities?
4. How are IT organizations approaching the evaluation and deployment of NV and/or SDN?
5. What is the role of organizations such as the ONF and the OpenDayLight consortium?
6. What approach are the key vendors taking relative to NV and SDN?
7. What should IT organizations do to get ready for NV and SDN?

The Guide will be published both in its entirety and in a serial fashion. This is the first of the serial publications. This publication will focus on NV. The three subsequent publications will focus on:

1. SDN
2. The Vendor Ecosystem
3. Planning for NV and SDN

In August and September of 2013 a survey was given to the subscribers of Webtorials. Throughout this document, the IT professionals who responded to the surveys will be referred to as *The Survey Respondents*.

Introduction

Over the last couple of years a number of approaches to NV have emerged that are focused on addressing the limitations of the traditional techniques for network virtualization (e.g., 802.1Q VLANs and Virtual Routing and Forwarding (VRFs)). All of these approaches are based on creating a number of virtual Layer 2 or Layer 3 networks that are supported by a common physical infrastructure. The basic idea is to virtualize the network in a manner analogous to compute server virtualization. As a result of these developments, network designers will have the opportunity to choose among the following NV alternatives.

1. Traditional NV
2. Overlay Network Virtualization via Tunneling
3. Software Defined NV via Flow Table Segmentation
4. A combination of the above alternatives

The Survey Respondents were asked to indicate how their organization defines network virtualization and multiple answers were allowed. The survey question focused on the emerging forms of network virtualization – bullets 2 and 3 in the preceding list. As indicated in Table 1, some of the the emerging forms of network virtualization are based on a device referred to as a controller. As is described below, one of the key roles of a controller is to serve as a central repository of address mappings.

The responses to this question are shown in **Table 1**.

Table 1: Characterization of NV Solutions

Definition of Network Virtualization	Percentage of Respondents
It is based on overlays using protocols such as VXLAN, NVGRE or STT but it does not involve a controller	21.0%
It is based on overlays and a controller. It may or may not use protocols such as VXLAN, NVGRE or STT	39.1%
It is part of a software defined network and may be based on segregating traffic flows	36.2%
Don't know	17.7%
Other	4.5%

The data in **Table 1** indicates that of the emerging forms of network virtualization, the controller-based approaches to NV are by a wide margin the most popular.

VXLAN, NVGRE and STT are all draft IETF standards. To understand the role that standards play in the selection of NV solutions, The Survey Respondents were asked how important it was to their organization that NV solutions are based on open standards. Their responses are shown in **Table 2**.

Table 2: Importance of Open Standards	
Level of Importance	**Percentage of Respondents**
Extremely important	**16.0%**
Very important	32.1%
Moderately important	24.7%
Somewhat important	14.4%
Not important	7.4%
Don't know	5.3%

The data in **Table 2** indicates that NV solutions that are build on open standards are either very or extremely important to roughly half of The Survey Respondents.

Traditional NV & The NV Use Case

One-to-many virtualization of network entities is not a new concept. The most common traditional applications of the virtualization concept to networks are VRF instances and VLANs.

VRF is a form of Layer 3 network virtualization in which a physical router supports multiple virtual router (VR) instances, each running its own routing protocol instance and maintaining its own forwarding table. Unlike VLANs, VRF does not use a tag in the packet header to designate the specific VRF to which a packet belongs. The appropriate VRF is derived at each hop based on the incoming interface and information in the frame. An additional requirement is that each intermediate router on the end-to-end path traversed by a packet needs to be configured with a VRF instance that can forward that packet.

VLANs partition the standard Ethernet network into as many as 4,096 broadcast domains as designated by a 12 bit VLAN ID tag in the Ethernet header. VLANs have been a convenient means of isolating different types of traffic that share a common switched LAN infrastructure. In data centers making extensive use of server virtualization, the limited number of available VLAN IDs can present problems, especially in cases where a large number of tenants need to be supported, each of whom requires multiple VLANs. In contrast to this limitation of VLANs, part of the use case for the NV approaches that are described in The Guide is that these approaches enable IT organizations to establish virtual Ethernet networks without being constrained to only having 4,096 VLAN IDs.

Server virtualization is another factor that is driving the adoption of the approaches to NV that are described in this sub-section of The Guide. Due to server virtualization, virtual machines (VMs) can be dynamically created and moved, both within a data center and between data centers. Extending VLANs across a data center via 802.1Q trunks to support VM mobility adds operational cost and complexity due to the fact that each switch in end-to-end path has to be manually reconfigured. In data centers based on Layer 2 server-to-server connectivity, large numbers of VMs, each with its own MAC address, can also place a burden on the forwarding tables capacities of Layer 2 switches. A major component of the value proposition for the NV approaches that are described in The Guide is that they support the dynamic movement, replication and allocation of virtual resources without manual intervention. Another component of the value proposition for these approaches is that they avoid the issue of needing more MAC addresses than data center LAN switches can typically support.

The value proposition of network overlay solutions is expanded upon in the following sub-section. As is also described below, one characteristic of NV solutions that IT organizations need to understand is whether the solution enables the dynamic movement of virtual resources within a data center; between data centers; or between a data center and a branch or campus facility. A related characteristic that IT organizations need to understand is whether the solution leverages standards based protocols to federate with other NV solutions.

Network Overlays via Tunneling: Benefits & Limitations

A number of approaches to network virtualization leverage tunneling and encapsulation techniques to construct multiple virtual network topologies overlaid on a common physical network. A virtual network (VN) can be a Layer 2 network or a Layer 3 network, while the physical network can be Layer 2, Layer 3 or a combination depending on the overlay technology. With overlays, the outer (encapsulating) header includes a field (generally up to 24 bits wide) that carries a virtual network instance ID (VNID) that specifies the virtual network designated to forward the packet.

Virtual network overlays can provide a wide range of benefits, including:

- Virtualization is performed at the network edge, while the remainder of the L2/L3 network remains unchanged and doesn't need any configuration change in order to support the virtualization of the network. The most common approach is to perform the encapsulation at the hypervisor vSwitch, which acts as the virtual tunnel endpoint (VTEP) or network virtualization edge (NVE). As a result, overlay NV solutions can generally be implemented over existing networks as either an enhancement to the conventional distributed network architecture, or as a step toward an SDN architecture.

- Support for essentially unlimited numbers of VNs as the 24 bits that are typically used by network overlays to identify VNs can identify slightly more than 16 million VN IDs. While theoretically NV solutions can support 16 million VNs, practical limits are often in the range of 16,000 to 32,000 VNs.

- Decoupling of the virtual network topology from the physical network Infrastructure and decoupling of the "virtual" MAC and/or IP addresses used by VMs from the infrastructure IP addresses used by the physical data center core network. The decoupling avoids issues such as limited MAC table size in physical switches.

- Support for VM mobility independent of the physical network. If a VM changes location, even to a new subnet in the physical network, the switches at the edge of the overlay simply update mapping tables to reflect the new physical location of the VM. The network for a new VM can be be provisioned entirely at the edge of the network.

- Ability to manage overlapping IP addresses between multiple tenants.

- Support for multi-path forwarding within virtual networks.

- Ease of provisioning virtual appliances in the data path. Network services resident on VMs can be chained together (a.k.a., service chaining) with point-and-click simplicity under the control of NV software.

- For controller-based NV solutions, the controller is not in the data path, and so it does not present a potential bottleneck.

The Survey Respondents were given a set of 15 possible challenges and opportunities and were asked to indicate which challenges and opportunities they thought that NV solutions could help them to respond to. The Survey Respondents were allowed to indicate multiple challenges and opportunities. The top 5 challenges and opportunities are shown in **Table 3.**

Table 3: Use Cases for NV Solutions	
Challenge/Opportunity	**Percentage of Respondents**
Better utilize network resources	44.0%
Support the dynamic movement, replication and allocation of virtual resources	39.1%
Establish virtual Ethernet networks without the limit and configuration burden of VLANs	32.5%
More easily scale network functionality	31.7%
Reduce OPEX	30.5%

Given the similarity of the second and third entries in **Table 3**, it follows that the primary value that IT organizations see in NV solutions is the ability to dynamically implement virtual Ethernet networks that can support the dynamic movement, replication and allocation of virtual resources.

Some of the limitations of overlay NV solutions include:

- Virtual and physical networks are separate entities, possibly with separate service assurance solutions, policy management, provisioning, and control points.
- As the virtual networks grow and evolve, the physical network does not automatically adapt to the changes. As a result, overlay NV requires a lightly oversubscribed or non-oversubscribed physical underlay network.
- Gateways between the virtual network and systems and network service points on the physical network may need to pass high volumes of traffic. If a software gateway running on a VM or a dedicated appliance has insufficient processing power, hardware support for the gateway functionality may be required in physical switches or network service appliances. Some of the more recent merchant silicon switching chips support gateway functionality for VXLAN which is the most popular encapsulation protocol.
- Some value-added features in existing networks cannot be leveraged due to encapsulation. For example, the physical network loses its ability to provide differentiated services based on the content of the packet header.

NV solutions also create some management challenges. For example, one of the primary benefits of overlay solutions is the ability to support multiple VNs running on top of the physical network. Effective operations management requires that IT organizations have tools that give them clear visibility into the relationships between virtual and physical networks and their component devices. When performance or availability problems occur, both root cause analysis and impact analysis require bilateral mapping between the physical and virtual infrastructures.

Both increasing and complicating the need for the visibility described in the preceding paragraph is the ability of NV solutions to do service chaining. The phrase *service chaining* refers to the ability to steer VM-VM traffic flows through a sequence of physical or virtual servers that provide network services, such as firewalls, IPS/IDS, DPI, or load balancers. The primary focus of service chaining is on services provided by virtual appliances. Most SDN or NV solutions

provide service chaining. For SDN, the controller configures the forwarding plane switches to direct the flows along the desired paths, For NV, the controller adjust the FIBs of the vSwitches/vRouters to force the traffic through the right sequence of VMs. Network Function Virtualization, discussed in the next section of The Guide, is basically service chaining that focuses on network services/functions provided by virtual appliances, but isn't necessarily dependent on SDN or NV.

The bottom line is that IT organizations need visibility not just into the overlay NV solution but into the complete solution and all of its components; e.g., firewalls, load balancers.

The Survey Respondents were given a set of 12 inhibitors to the adoption of NV and were asked to indicate the two biggest inhibitors to their company adopting NV sometime in the next two years. The top 5 inhibitors are shown in **Table 4**.

Table 4: Inhibitors to the Adoption of NV Solutions

Inhibitor	% of Respondents
The immaturity of the current products	29.6%
The lack of resources to evaluate NV	29.2%
Other technology and/or business priorities	28.8%
The immaturity of the enabling technologies	29.6%
The confusion and lack of definition in terms of vendors' strategies	18.1%

One interesting observation that can be drawn from the data in **Table 4** is that IT organizations are not avoiding implementing NV solutions because they don't see value in them. Rather, the key factors inhibiting the adoption of NV solutions are the same factors that typically inhibit the adoption of any new technology or way of implementing technology: Immaturity of products and strategies; confusion; and lack of resources.

The Survey Respondents were asked to indicate the impact they thought that NV would have on security and network management. Their responses are shown in **Table 5** and **Table 6**.

Table 5: Impact of NV on Security

Impact on Security	% of Respondents
Networks will be much more secure	6.2%
Networks will be somewhat more secure	33.7%
NV will have no impact on network security	23.5%
Networks will be somewhat less secure	14.0%
Networks will be much less secure	2.5%
Don't know	20.2%

Table 6: Impact of NV on Management	
Impact on Management	**% of Respondents**
Networks will be much easier to manage	21.8%
Networks will be somewhat easier to manage	52.3%
NV will have no impact on management	4.5%
Networks will be somewhat more difficult to manage	9.9%
Networks will be much more difficult to manage	4.5%
Don't know	7.0%

One conclusion that can be drawn from the data in **Table 5** and **Table 6** is that The Survey Respondents generally think that implementing NV solutions will make their networks more secure and easier to manage. As such, security and ease of management can potentially be looked at as benefits of implementing NV solutions.

Cloud Orchestration

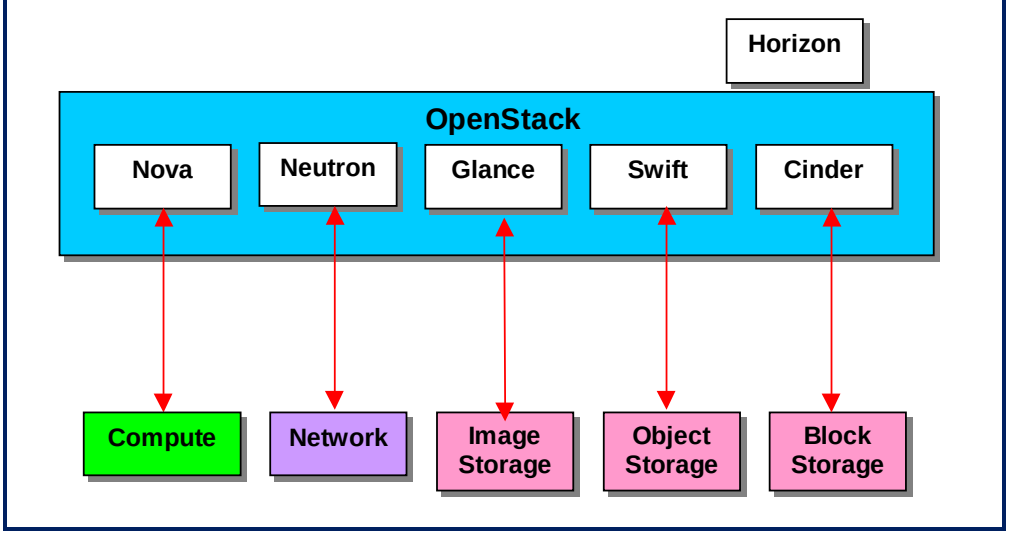
Cloud Orchestration platforms have evolved as a means of automating and facilitating the process of configuring pools of data center resources in order to provide a range of cloud or cloud-like services, such as Infrastructure as a Service (IaaS) solutions. The Orchestrator's role is to manipulate the basic resources of the data center (i.e., VMs, networks, storage, and applications) at a very high level of abstraction to create the service. Orchestration is most effective when the data center is fully virtualized, facilitating software control/reconfiguration and automation. As a result, there is a naturally affinity between Orchestration and software-based network controllers, such as NV controllers or SDN controllers.

OpenStack is a cloud computing orchestration project offering free open source software released under the terms of the Apache License. The project is managed by the OpenStack Foundation, a non-profit corporate entity established in September 2012 to promote OpenStack software and its community. Apache CloudStack is another open source Apache Licensed orchestration system. Eucalyptus is a third open source orchestrator with tight technical ties to Amazon Web Services (AWS).

In addition, there are a number of proprietary orchestrators that offer open APIs to allow integration across vendor boundaries. These include VMware's vCloud Director and IBM's SmartCloud Orchestrator.

Figure 1 shows a block diagram of the OpenStack system, including the OpenStack modules that are used to control resource pools in the data center. Horizon is the OpenStack Dashboard that provides administrators and users a graphical interface to access, provision and automate cloud-based resources.

Figure 1: OpenStack



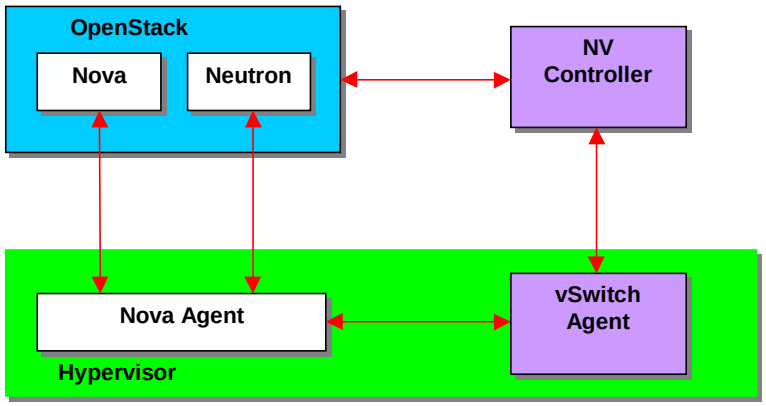
Neutron (formerly called Quantum) allows users to create their own networks, provide connectivity for servers and devices, and control traffic. With appropriate Neutron plug-ins, administrators can take advantage of various NV and SDN solutions to allow for multi-tenancy and scalability. OpenStack networking also has an extension framework allowing additional network services, such as intrusion detection systems (IDS), load balancing, firewalls and virtual private networks (VPN) to be deployed and managed.

In conjunction with the Orchestrator, the role of the SDN or NV controller is to translate the abstract model created on the Orchestrator into the appropriate configuration of the virtual and physical resources that will deliver the desired service. For example, the orchestrator can instruct the controller to perform a variety of workflows, including:

- Create a VM
- Assign a VM to a Virtual Network (VN)
- Connect a VM to an external network
- Apply a security policy to a group of VMs or a Virtual Network
- Attach Network Services to a VM or chain Network Services between VMs

Figure 2 provides a high level depiction of how an orchestrator (OpenStack) and a NV controller might interact to place a VM into service within a VN.

Figure 2: VM Creation Workflow with OpenStack

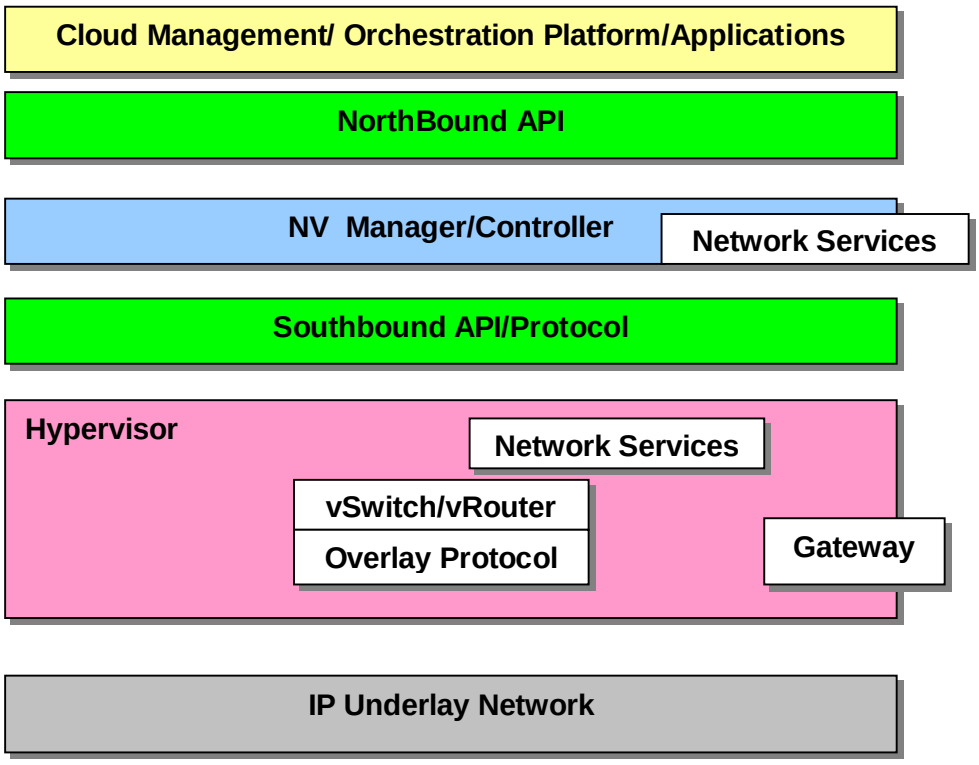


The Nova module in OpenStack instructs the Nova Agent in the hypervisor to create the VM. The Nova agent communicates with the Neutron module in OpenStack to learn the network attributes of the VM. The Nova agent then informs the vSwitch agent to configure the virtual network for the VM and then the controller provides the route table entries needed by the vSwitch.

Controller Based NV Solution Architecture

A Network Virtualization Solution typically has an architecture similar to the one shown in **Figure 3**. The main components are typically the NV Controller, hypervisor-resident vSwitches/vRouters, and gateways that provide connectivity from virtual networks to traditional network segments; e.g., VLANs, non-virtualized servers, or Internet routers. The controller function is generally supported by a high availability (HA) cluster or another HA configuration. Controller functionality may be comprised of a number of sub-functions running on different servers. Cloud Management/Orchestration is typically obtained from a third party and network services may be integrated with the controller, integrated via virtual appliances, or possibly integrated via physical appliances through the gateway.

Figure 3: Network Virtualization Architecture



Criteria to Evaluate Overlay NV Solutions

One of the primary criterion that IT organizations should use relative to evaluating overlay network virtualization solutions is how well it solves the problem(s) that the IT organization is looking to solve. For example, can the solution enable the IT organization to move workloads between data centers? Between a data center and a branch office?

Other solution level criteria that IT organizations should evaluate include:

- Does the solution federate and hence interoperate with other solutions?
- What interaction, if any, is there between the virtual networks and the physical networks?
- What management functionality is provided into both the virtual and physical networks?
- Does the solution support service chaining?

The main technical differences between the various overlay NV solutions that IT organizations should evaluate fall into the following categories:

- **Encapsulation formats**. Some of the tunneling/encapsulation protocols that provide network virtualization of the data center include VXLAN, NVGRE, STT, and SPB MAC-in-MAC (SPBM). Both the IEEE and the IETF have already standardized SPB. It is unclear as to whether or not all of the other proposals will become standards.
- **Tunnel control plane functionality** that allows ingress (encapsulating) devices to map a frame to the appropriate egress (decapsulating) device. The first-hop overlay device implements a mapping operation that determines where the encapsulated packet should be sent to reach its intended destination VM. Specifically, the mapping function maps the destination address (either L2 or L3) of a packet received from a VM into the corresponding destination address of the egress NVE device. The main differences here are whether a controller is used and the functionality of the controller.

  Some of the initial, controller-less approaches to network virtualization relied on IP multicast as a way to disseminate address mappings. A more common solution is based on a central repository of address mappings housed in a controller. Vendors frequently refer to controller-based overlay NV solutions as SDN, while a more descriptive terminology might be Software Defined Overlay Network Virtualization.
- **vSwitches supported.** A number of vSwitches are based to some degree on the open source Open vSwitch (OVS)[1], while other vSwitches are of proprietary design. Another point of differentiation is whether the vSwitch is a virtual router as well as being an encapsulating Layer 2 switch. With Layer 3 functionality, a vSwitch can forward traffic between VMs on the same hypervisor that are in different subnets and can be used to implement Layer 3 VNs. Where the tunneling vSwitch has full Layer 3 functionality, the majority of intelligence can be implemented at the edge of network, allowing the underlay network to be implemented as a simple Layer 2 fabric.

[1] While based on OVS, many vSwitches have implemented proprietary extensions to OVS.

- **Broadcast/Multicast delivery** within a given virtual network. NVEs need a way to deliver multi-destination packets to other NVEs with destination VMs. There are three different approaches that can be taken:

    - The multicast capabilities of the underlay network can be used
    - The NVEs can replicate the packets and unicast a copy across the underlay network to each NVE currently participating in the VN.
    - The NVE can send the packet to a distribution server which replicates and unicasts the packets on the behalf of the NVEs.

- **Protocols.** Another characteristic of centralized controller solutions is the choice of Southbound protocols/APIs employed between the NV controller and the NVE and the choice of Northbound protocols/APIs used between the NV controller and cloud management systems and hypervisor management systems. If the southbound protocols are standardized, the NVE can potentially communicate with different types of NV controllers or controllers from different vendors. Some the alternatives here include OpenFlow, BGP, and CLI shell scripts.

  If the northbound protocols are standardized, the controller can be integrated with network services from ISVs or different types of third party orchestration systems. Most overlay NV controllers support a RESTful Web API for integration with cloud management and orchestration systems. With both southbound and northbound APIs the most important question becomes which third party switches, applications, virtual appliances, and orchestration systems have been certified and are supported by the overlay NV vendor.

- **VN Extension over the WAN.** VN extension over the WAN can generally be accomplished with most NV solutions. However, in some cases the encapsulation used over the wide area may differ from that used within the data center. Some of the encapsulation techniques used for VN extension over the WAN include MPLS VPNs and two proprietary protocols from Cisco: Overlay Transport Virtualization (OTV) and Locator/ID Separation Protocol (LISP). OTV is optimized for inter-data center VLAN extension over the WAN or Internet using MAC-in-IP encapsulation. It prevents flooding of unknown destinations across the WAN by advertising MAC address reachability using IS-IS routing protocol extensions. LISP is an encapsulating IP-in-IP technology that allows end systems to keep their IP address (ID) even as they move to a different subnet within the network (Location). By using LISP VM-Mobility, IP endpoints such as VMs can be relocated anywhere regardless of their IP addresses while maintaining direct path routing of client traffic. LISP also supports multi-tenant environments with Layer 3 virtual networks created by mapping VRFs to LISP instance-IDs. Inter-data center network virtualization could also potentially be based on Layer 3 vSwitches that support MPLS VPNS and implement network virtualization using RFC 4023 MPLS over IP/GRE tunnels through an IP enterprise network to connect to an MPLS VPN service. SPBM is unique in that it offers extensions over the WAN natively without requiring additional protocols such as OTV or MPLS VPNs.

The remainder of this sub-section of The Guide focuses on the primary differentiating features of Overlay NV solutions: tunnel encapsulation and tunnel control.

Tunnel Encapsulation

VXLAN: Virtual eXtensible LAN (VXLAN)[2] virtualizes the network by creating a Layer 2 overlay on a Layer 3 network via MAC-in-UDP encapsulation. The VXLAN segment is a Layer 3 construct that replaces the VLAN as the mechanism that segments the data center LAN for VMs. Therefore, a VM can only communicate or migrate within a VXLAN segment. The VXLAN segment has a 24-bit VXLAN Network identifier. VXLAN is transparent to the VM, which still communicates using MAC addresses. The VXLAN encapsulation is performed through a function known as the VXLAN Tunnel End Point (VTEP), typically a hypervisor vSwitch or a possibly a physical access switch. The encapsulation allows Layer 2 communications with any end points that are within the same VXLAN segment even if these end points are in a different IP subnet. This allows live migrations to transcend Layer 3 boundaries. Since MAC frames are encapsulated within IP packets, there is no need for the individual Layer 2 physical switches to learn MAC addresses. This alleviates MAC table hardware capacity issues on these switches. Overlapping IP and MAC addresses are handled by the VXLAN ID, which acts as a qualifier/identifier for the specific VXLAN segment within which those addresses are valid.

As noted, VXLANs use a MAC-in-UDP encapsulation. One of the reasons for this is that modern Layer 3 devices parse the 5-tuple (including Layer 4 source and destination ports). While VXLAN uses a well-known destination UDP port, the source UDP port can be any value. As a result, a VTEP can spread all the flows from a single VM across many UDP source ports. This allows for efficient load balancing across link aggregation groups (LAGs) and intermediate multi-pathing fabrics even in the case of multiple flows between just two VMs.

Where VXLAN nodes on a VXLAN overlay network need to communicate with nodes on a legacy (i.e., VLAN) portion of the network, a VXLAN gateway can be used to perform the required tunnel termination functions including encapsulation/decapsulation. The gateway functionality could be implemented in either hardware or software.

VXLAN is supported by a number of vendors including Cisco Systems, VMware, IBM, and Nuage Networks. Avaya's SPBM implementation (Fabric Connect) can also support a VXLAN deployment, acting as a transport layer providing optimized IP Routing and Multicast for VXLAN-attached services.

STT: Stateless Transport Tunneling (STT)[3] is a second overlay technology for creating Layer 2 virtual networks over a Layer 2/3 physical network within the data center. Conceptually, there are a number of similarities between VXLAN and STT. The tunnel endpoints are typically provided by hypervisor vSwitches, the VNID is 24 bits wide, and the transport source header is manipulated to take advantage of multipathing. STT encapsulation differs from VXLAN in two ways. First, it uses a stateless TCP-like header inside the IP header that allows tunnel endpoints within end systems to take advantage of TCP segmentation offload (TSO) capabilities of existing TOE server NICs. The benefits to the host include lower CPU utilization and higher utilization of 10 Gigabit Ethernet access links. STT generates a source port number based on hashing the header fields of the inner packet to ensure efficient load balancing over LAGs and multi-pathing fabrics. STT also allocates more header space to the per-packet metadata, which

[2] http://searchservervirtualization.techtarget.com/news/2240074318/VMware-Cisco-propose-VXLAN-for-VM-mobility

[3] http://tools.ietf.org/html/draft-davie-stt-01

provides added flexibility for the virtual network tunnel control plane. With these features, STT is optimized for hypervisor vSwitches as the encapsulation/decapsulation tunnel endpoints. The initial implementations of Network Virtualization using STT from Nicira Networks are based on OpenFlow-like hypervisor vSwitches (Open vSwitches) and a centralized control plane for tunnel management via downloading mapping tables to the vSwitches.

NVGRE: Network Virtualization using Generic Router Encapsulation (NVGRE)[4] uses the GRE tunneling protocol defined by RFC 2784 and RFC 2890. NVGRE is similar in most respects to VXLAN with two major exceptions. While GRE encapsulation is not new, most network devices do not parse GRE headers in hardware, which may lead to performance issues and issues with 5-tuple hashes for traffic distribution in multi-path data center LANs. With GRE hashing generally involves the GRE key. One initial implementation of NVGRE from Microsoft relies on Layer 3 vSwitches whose mapping tables and routing tables are downloaded from the vSwitch manager. Downloads are performed via a command-line shell and associated scripting language.

SPBM[5]: IEEE 802.1aq/IETF 6329 Shortest Path Bridging MAC-in-MAC uses IEEE 802.1ah MAC-in-MAC encapsulation and the IS-IS routing protocol to provide Layer 2 network virtualization and VLAN extension in addition to a loop-free equal cost multi-path Layer 2 forwarding functionality. VLAN extension is enabled by the 24-bit Service IDs (I-SIDs) that are part of the outer MAC encapsulation. Unlike other network virtualization solutions, no changes are required in the hypervisor vSwitches or NICs and switching hardware already exists that supports IEEE 802.1ah MAC-in-MAC encapsulation. For SPBM, the control plane is provided by the IS-IS routing protocol.

SPBM can also be extended to support Layer 3 forwarding and Layer 3 virtualization as described in the IP/SPB IETF draft using IP encapsulated in the outer SPBM header. This specification identifies how SPBM nodes can perform Inter-ISID or inter-VLAN routing. IP/SPB also provides for Layer 3 VSNs by extending VRF instances at the edge of the network across the SPBM network without requiring that the core switches also support VRF instances. VLAN-extensions and VRF-extensions can run in parallel on the same SPB network to provide isolation of both Layer 2 and Layer 3 traffic for multi-tenant environments. With SPBM, only those Switches that define the SPBM boundary need to be SPBM-capable. Switches not directly involved in mapping services to SPB service IDs don't require special hardware or software capabilities. SPBM isn't based on special vSwitches, data/control plane separation, or centralized controllers. SPBM hardware Switches are currently available from several vendors, including Avaya and Alcatel-Lucent.

Tunnel Control

As previously mentioned, initial implementations of VXLAN by Cisco and VMware use flooding as a distributed control solution based on Any Source Multicast (ASM) to disseminate end system location information. Because flooding requires processing by all the vSwitches in the multicast group, this type of control solution will not scale to support very large networks.
.
A more recent approach is to implement tunnel control as a centralized controller function. A control plane protocol that carries both MAC and IP addresses can eliminate the need for ARP.

[4] http://datatracker.ietf.org/doc/draft-sridharan-virtualization-nvgre/
[5] http://tools.ietf.org/html/draft-allan-l2vpn-spbm-evpn-00

One controller-based solution for VXLAN control, championed by IBM's Distributed Overlay Virtual Ethernet (DOVE) initiative, is to use a DNS-like network service to map the VM's IP address to the egress VTEP's IP address. IBM's solution does not require Multi Cast enablement in the physical network. IBM's Controller based solution has built-in IP routing capability.

In another controller-based approach, used by Nicira Networks, the controller maintains a data base of Open vSwitches (OVS) in the network and proactively updates OVS mapping tables via OpenFlow to create new tunnels when VMs are created or moved. The Nicira controller focuses on the virtual network topology and is oblivious to the topology of the core physical network. The controller is integrated with hypervisor and cloud management systems to learn of changes in the population of VMs.

A third controller approach, used by Nuage Networks and Netsocket, involves the controller maintaining a full topology of the virtual and physical network and maintaining the full address mapping and routing tables derived from standard routing protocols, such as OSPF, IS-IS, or BGP. The portion of the table needed by the vSwitch is disseminated from the controller to the vSwitches via the OpenFlow protocol. The Nuage Networks' vSwitches use VXLAN to encapsulate L2 traffic and GRE to encapsulate L3 traffic.

Comparison of Network Virtualization Solutions

The following table (**Table 7**) provides a high level summary of the primary features of some of the Network Virtualization solutions that are available or have been recently announced. Note that the solutions described in columns two and three (Cisco, VMware) are not based on a controller.

Table 7: Network Virtualization features

	Cisco	VMware	IBM	VMware/ Nicira	Nuage Networks	Avaya	Netsocket	Juniper
Product	**Nexus 1000v**	**VSphere DS**	**SDN-VE**	**NSX**	**VSP**	**Fabric Connect**	**NVN**	**Contrail**
Overlay	VXLAN	VXLAN	VXLAN	VXLAN STT?	VXLAN	SPBM	GRE	MPLS/GRE MPLS/UDP VXLAN
VM-NVE Address Learning	VTEP Multicast flooding	VTEP Multicast flooding	Pull From Controller's Directory	Push From Controller's Data Base	Push From Controller's Map Table	IS-IS SPB on physical switch	Push From Controller's Map Table	Push From Controller's Map Table
Broadcast / Multicast within VN	via underlay Multicast	via underlay Multicast	distribution server replication	distribution server replication	dVRS packet replication	via SPB multicast		VRouter packet replication or proxy
Controller Topology Awareness	na	na	Virtual Networks	Virtual Networks	Entire Network	Entire Network	Entire Network	Entire Network
Controller to NVE Protocol	NX-OS CLI	VMware API	Open source submitted to OpenDaylight	OpenFlow NSX API	OpenFlow	IS-IS	vFlow or OpenFlow	XMPP
vSwitch	Nexus 1000v	VDS	SDN-VE vSwitch	VDS, Open vSwitch**	dVRS (Open vSwitch**)	Native to Hypervisor	vFlowSwitch	v Contrail vRouter
vSwitch L3	no	no	yes	yes	yes	na	yes	yes
Gateway Support in Physical Switches				Arista 7150s Brocade ADX	Nuage Networks 7850 VSG	na		
Hypervisors	ESXi, Hyper-V, XEN, KVM	ESXi	ESXi KVM	vSphere. ESXi, XEN, KVM	ESXi, Hyper-V, XEN, KVM	ESXi, Hyper-V, XEN, KVM	Hyper-V ESXi Xen, KVM	KVM, XEN
Controller Federation					via MP-BGP			BGP
DC-DC encapsulation	OTV	OTV	VXLAN	GRE	MPLS over GRE to PE router	Over an SPBM WAN	GRE	MPLS/GRE
	OpenStack vCloud	OpenStack vCloud	OpenStack	OpenStack CloudStack vCloud	OpenStack CloudStack vCloud	OpenStack Integration in controller	OpenStack System Ctr.	OpenStack.

na = not applicable ***** = with proprietary extensions***

Software Defined NV via Flow Table Segmentation

Network virtualization can also be implemented as an application that runs on an SDN controller. Virtual networks are defined by policies that map flows to the appropriate virtual network based on L1-L4 portions of the header. With this type of SDN-based NV, there is no need for tunnels and encapsulation protocols. One example of an NV application is the Big Virtual Switch that runs on the Big Network Controller from Big Switch Networks. The Big Network Controller implements VNs by configuring forwarding tables in OpenFlow physical and virtual switches. The OpenFlow switches can be from a variety of traditional switch vendors. Another alternative is to use Big Switch Switch Light OpenFlow thin software agent running on bare metal Ethernet switches based on Broadcom merchant silicon or on virtual switches.

By exploiting the capability of OpenFlow to deal with encapsulation and de-encapsulation, the SDN controller NV application can also be used to implement overlay VNs running over a conventional L2/L3 network, or a hybrid network based partially on pure SDN VNs and partially on SDN NVs with OpenFLow virtual switches and a conventional core network.

Another slightly different approach to an NV application for SDN controllers is the Virtual Tenant Network (VTN) application developed by NEC and recently accepted as an application by the OpenDaylight consortium. The VTN solution provides a layer of abstraction between the virtual network and the physical network. In the event of a failed link, the VTN can detect and redirect the affected flows within milliseconds. This avoids the re-convergence delay associated with traditional network protocols. The VTN also supports redirection, which enables use cases related to traffic steering and service chaining. In addition, the VTN physical control of the network supports flow based traffic engineering as well as 8-way ECMP.

VTN is based on a logical abstraction that decouples the VTN from the physical network. A virtual network can be designed and deployed using the following set of logical network elements:

- vBridge L2 switch function.
- vRouter router function.
- vTEP virtual Tunnel End Point.
- vTunnel Tunnel.
- vBypass connectivity between controlled networks.
- vInterface end point on the virtual node.
- vLink L1 connectivity between virtual interfaces.

Using these elements allows the user can define a logical network with the look and feel of conventional L2/L3 network. VTN can also by used to implement an overlay network, an OpenFlow network, or a hybrid overlay/OpenFlow network. Once the network is designed on VTN, it can automatically be mapped onto the underlying physical network, and configured on the individual switches leveraging an SDN control protocol, Typically this would be OpenFlow. Mapping is used to identify the VTN to which each packet transmitted or received by an OpenFlow switch belongs, as well as which interfaces on the OpenFlow switch can transmit or receive that packet. Flows are mapped to a VTN vBridge based on the ingress port on the OpenFlow switch, the source MAC address or the VLAN ID.

Enterprise Plans for NV Adoption

The Survey Respondents were asked a series of questions about their current position relative to evaluating and adopting NV solutions and how that position might change over the next two to three years. In the first of those questions, The Survey Respondents were given a set of alternatives and were asked to indicate the alternatives that described their company's current approach to implementing NV solutions. Their responses are shown in **Table 8**.

Table 8: Current Approaches to Adopting NV Solutions

Approach to Adoption NV Solutions	% of Respondents
We have not made any analysis of NV	25.5%
We will likely analyze NV sometime in the next year	25.5%
We are currently actively analyzing the potential value that NV offers	24.7%
We expect that within a year that we will be running NV either in a lab or in a limited trial	13.6%
We are currently actively analyzing vendors' NV strategies and offerings	11.5%
We currently are running NV either in a lab or in a limited trial	9.9%
We currently are running NV somewhere in our production network	7.4%
We looked at NV and decided to not do anything with NV over the next year	6.2%
We expect that within a year that we will be running NV somewhere in our production network	5.8%
Don't know	4.9%

The data in **Table 8** indicates that while there is currently little deployment of NV, there is a lot of activity and interest relative to analyzing NV solutions. The data in Table 8 also suggests that over the next year the percentage of IT organizations that are either running NV somewhere in their production network, or in a lab or limited trial, will double.

The Survey Respondents were given a two-year time frame and were asked to indicate where in their infrastructure their organization was likely to implement NV solutions. (Multiple responses were allowed) Their responses are shown in **Table 9**.

Table 9: Likely Deployment of NV Solutions

Focus of Future NV Implementation	% of Respondents
Data Center	58.0%
Branch and/or Campus	25.1%
WAN	18.5%
We are unlikely to implement NV in the next two years	15.6%
Don't know	10.7%
We are likely to acquire a WAN service that is based on NV	9.5%

The data in **Table 9** indicates that IT organizations will primarily implement NV solutions within a data center. However, the data also indicates that a sizeable percentage of IT organizations want to extend their NV solutions over the WAN and to also implement NV solutions in their branch and campus networks.

In the final question about their potential future use of NV solutions, The Survey Respondents were asked to indicate how broadly their data center networks will be based on NV three years from now. Their responses are shown in **Table 10**.

Table 10: Data Center Design in Three Years

Balance of NV and Traditional Approach	% of Respondents
Exclusively based on NV	3.3%
Mostly based on NV	25.1%
NV and traditional networking coexisting about equally	37.9%
Mostly traditional	16.9%
Exclusively traditional	4.1%
Don't know	12.8%

The data in **Table 10** indicates that the vast majority of The Survey Respondents expect that in three years that at least half of their data center networks will be based on NV.

About the Webtorials® Editorial/Analyst Division

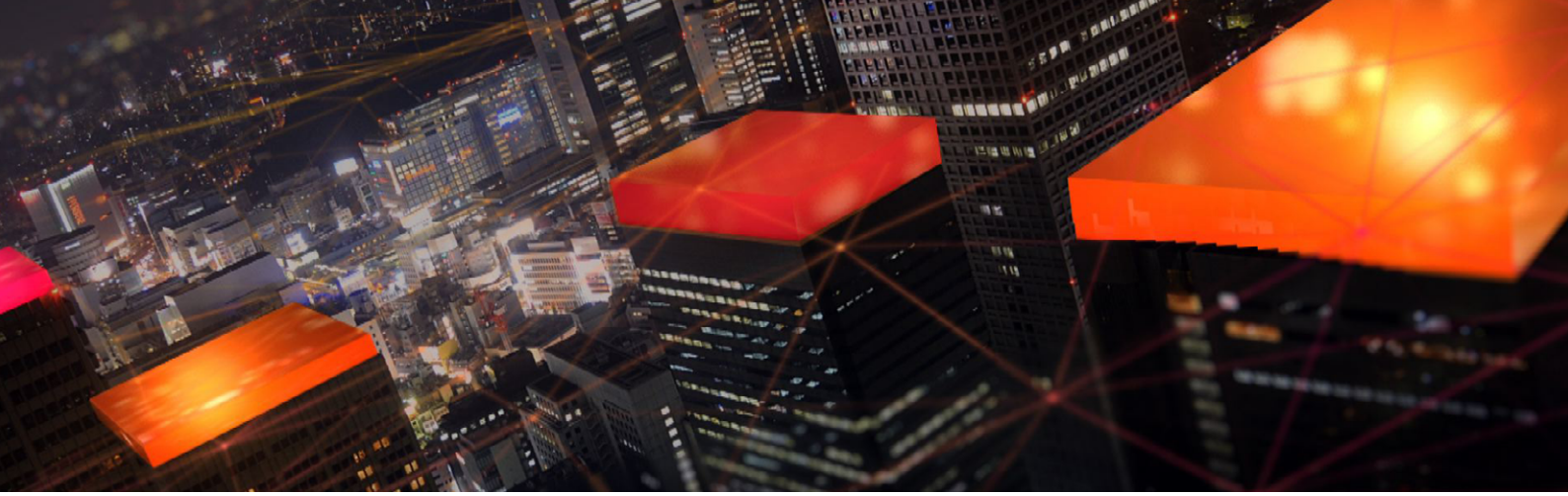
The Webtorials® Editorial/Analyst Division, a joint venture of industry veterans Steven Taylor and Jim Metzler, is devoted to performing in-depth analysis and research in focused areas such as Metro Ethernet and MPLS, as well as in areas that cross the traditional functional boundaries of IT, such as Unified Communications and Application Delivery. The Editorial/Analyst Division's focus is on providing actionable insight through custom research with a forward looking viewpoint. Through reports that examine industry dynamics from both a demand and a supply perspective, the firm educates the marketplace both on emerging trends and the role that IT products, services and processes play in responding to those trends.

Jim Metzler has a broad background in the IT industry. This includes being a software engineer, an engineering manager for high-speed data services for a major network service provider, a product manager for network hardware, a network manager at two Fortune 500 companies, and the principal of a consulting organization. In addition, he has created software tools for designing customer networks for a major network service provider and directed and performed market research at a major industry analyst firm. Jim's current interests include cloud networking and application delivery.

For more information and for additional Webtorials® Editorial/Analyst Division products, please contact Jim Metzler at *jim@webtorials.com* or Steven Taylor at *taylor@webtorials.com*.

Published by Webtorials Editorial/Analyst Division
www.Webtorials.com

Division Cofounders:
Jim Metzler
jim@webtorials.com
Steven Taylor
taylor@webtorials.com



The Consumable Datacenter Network

Taking cloud computing to the next level

The move to cloud computing and storage has changed the way Enterprise users access and consume data. Unfortunately, today's data communications networks aren't keeping pace with this dynamic business environment, and they're struggling to deliver consistent, on-demand connectivity.

That's where we come in. Nuage Networks™ closes the gap between the network and the cloud-based consumption model, creating an infrastructure in which network resources are as readily consumable as compute and storage resources. Our approach enables enterprises to transform the way they build and use their networks, which has a profound effect inside and across multiple datacenters. The transformation is also felt at the critical remote working environment, through a seamless connection to the Enterprise's Wide Area Network.

WOULDN'T IT BE NICE IF...

- Datacenter infrastructures were so simple and standards-based that you could break the vendor lock and work with whichever suppliers offered you the best solutions for your business?
- The network could expand and evolve transparently with the needs of applications, bypassing the datacenter's arbitrary boundaries?
- The datacenter network team could set up controlled, secure templates that application teams could use to deploy applications on the network for and by themselves — without manual transactions or unnecessary project overhead?

Before the move to the cloud, enterprises had to purchase large compute systems to meet the peak processing needs of a limited set of specific events, such as financial milestones (month end or year end), or annual retail events (holiday shopping). Outside of the specific events, the systems were underutilized. This approach was therefore expensive, both in terms of CAPEX and OPEX, requiring significant outlay for power, space and air-conditioning.

Cloud-based datacenters have unshackled the IT environment, making it possible for applications to request additional compute and storage on an as-needed basis. Peak demands can be provisioned "just in time", which lowers operational costs and provides the ability to share compute resources across applications.

The term "cloud" means many things to many people. We focus on two key benefits that cloud computing delivers to Enterprises:

Abstraction of the application from the infrastructure. Cloud computing separates the application from the physical compute and storage infrastructure. This allows workloads to be consistently configured remotely, and templated for mass deployment. End users don't need to worry about the location and specifications of individual hosts. Virtualization and cloud management tools abstract those details to make the infrastructure more readily consumable.

Customer self-fulfillment. Cloud Management Systems (CMS) like Alcatel-Lucent CloudBand™ and the abstraction layer enabled by server virtualization allow IT departments to minimize the tedious and cumbersome processing of application-to-network transactions. For example, IT can provision end customer access policies in the CMS to govern who is authorized to create virtual machine instances, in which location, how many are allowed, and who is the funding department. Users and work groups get instant application deployment, which in turn, makes the business more agile and responsive — critical

attributes in today's enterprise environment. At the same time, operational expenses associated with the handling of work orders is greatly reduced.

As a result of these innovations, Enterprises enjoy a powerful new IT environment in which applications can consume compute resources easily. However as the dynamic nature of cloud computing becomes mainstream, the underlying datacenter network is struggling to match the flexibility of the applications. In fact, most often the network is the weak link, inhibiting the enterprise's ability to profit from the benefits that moving to the cloud should provide.

While virtual compute resources can be instantiated in seconds, it often takes days for network connectivity to be configured and established. Furthermore, the static configurations used by today's networks do not provide the efficiencies and flexibility needed to drive maximum server utilization and application availability.

Consuming the Network

Nuage Networks ensures your network elements are as efficient and flexible as your cloud computing. The result is a choreographed datacenter environment where the compute resources and network work seamlessly.

Imagine the possibilities when network resources are easily consumable. A Nuage Networks datacenter network is as dynamic, automated and virtualized as the server infrastructure, and supports the needs of applications with instantaneous network connectivity.

Nuage Networks eliminates the constraints that have been limiting the datacenter network as it scales out to meet growing demand. With Nuage Networks, you can:

- Define the network service design per application
- Optimize your workload placement across datacenter zones or even across geo-diverse datacenters
- Maximize efficiency of your compute and storage resources

Nuage Networks paves the way for datacenters of the future to be the heartbeat of a powerful cloud infrastructure. Enterprises and user groups could conceive and consume their own secure slices of a robust multi-tenant infrastructure, with appropriate operational visibility and control.

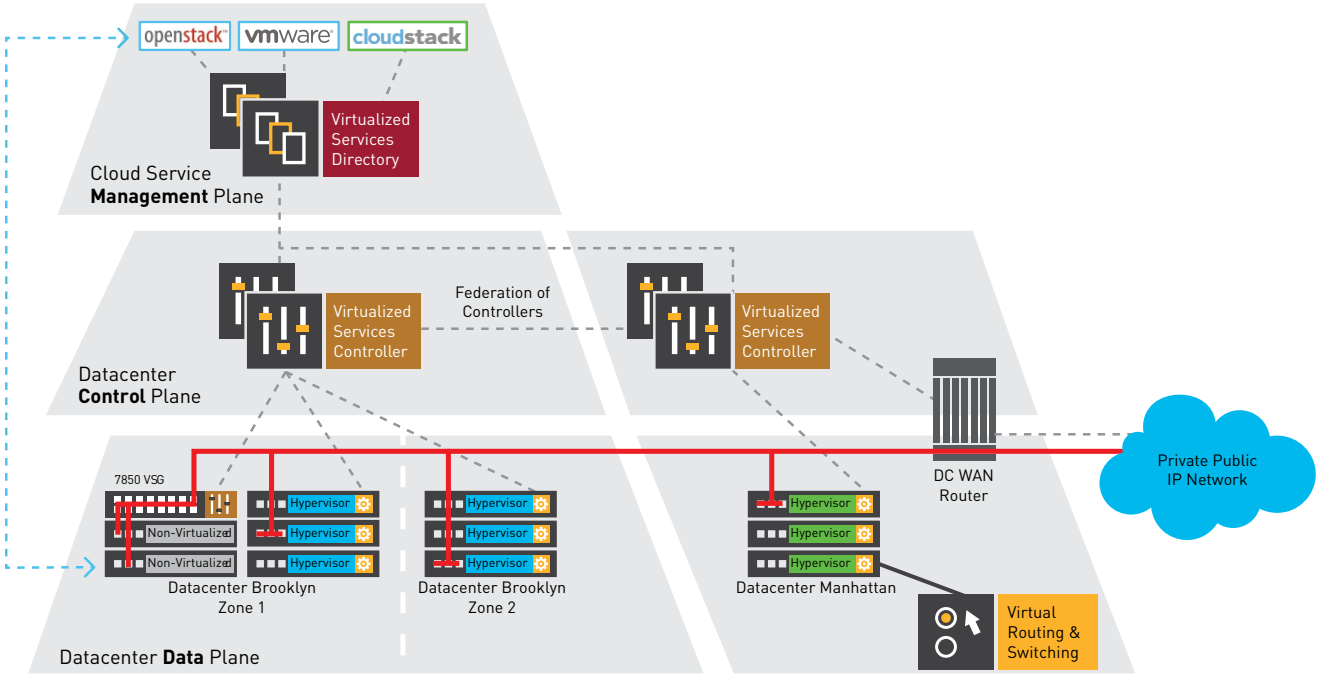
Nuage Networks Virtualized Services Platform

Nuage Networks Virtualized Services Platform (VSP) is the first network virtualization platform that addresses modern datacenter requirements for multi-tenancy, full-featured routing and security at scale. It also integrates seamlessly with wide area business VPN services. It is a software solution that transforms the physical network into a simple to manage, rack-once and wire-once, vendor-independent IP backplane. As a result, network resources within and across datacenters can be treated as an elastic resource pool of capacity that can be consumed and repurposed on demand. Nuage Networks enables unconstrained datacenter networks for the cloud era.

Nuage Networks delivers virtualization and automation of business networks through the three key elements in the Nuage Networks VSP:

Virtualized Services Directory (VSD). Configuration of networks is complex. To eliminate unnecessary complexity while leaving full control and visibility of applications with the IT administrator, the VSD abstracts networking constructs down to their base primitives in four categories: Connectivity Domains, Security, Quality of Service, and Analytics. This allows the requirements for network services to be expressed simply,

FIGURE 1. Nuage Networks Virtualized Services Platform



consistently, and in a repeatable manner. The critical need for mobility is also addressed, ensuring network services adjust gracefully and instantly as application endpoints and workloads move from virtual machines within or across datacenters.

The VSD also provides a rich permission-based multi-tenant interface to enable end user provisioning by application owners. Through its role-based hierarchy of permissions, the VSD eliminates operational delays and minimizes transactions between organizations while providing visibility and control of the network "slices" that each group is given in support of their application requirements.

Virtualized Services Controller (VSC). The VSC is an advanced SDN controller that manages the provisioning of virtual network services by programming the edges of the network using OpenFlow™. The VSC ensures that the network follows the application instantaneously. Parting with cumbersome and error-prone device-by-device manual provisioning, Nuage Networks introduces an event-triggered and pull-based configuration model. Once application events such as moves, adds or changes are detected, appropriate policy-based configurations are instantaneously applied. Leveraging Alcatel-Lucent's proven Service Router Operating System, which has been deployed in over 400 service provider networks worldwide for over a decade, the VSC runs a full and robust IP routing stack that allows it to communicate and seamlessly integrate into existing networks.

Virtual Routing and Switching (VRS) is a true hypervisor for the network. The first of its kind in the industry, the VRS fully virtualizes network offerings ranging from distributed virtual Layer 2, Layer 3 forwarding and Layer 4 security. These virtual network services leverage the existing network infrastructure and are offered in a standards-based manner compliant with IETF NVO3. Operators can use whatever servers, hypervisors, and cloud management systems they choose; the Nuage Networks solution abstracts and automates the cloud-networking infrastructure.

In many real-world installations, datacenter environments are a mix of virtualized and non-virtualized assets. To help all datacenters benefit from automation and network virtualization, Nuage Networks supports the full range of options. Software gateways such as the Nuage VRS-G are ideal for environments with relatively low density of bare metal servers and appliances, just as hardware VTEPs from our ecosystem partners provide a viable alternative for certain use cases and environments. For environments with significant investment in bare metal servers and appliances, a new breed of high performance gateway is needed.

The **Nuage Networks 7850 Virtualized Services Gateway (VSG)** is a high-performance gateway that extends Nuage Networks SDN 2.0 functionality seamlessly between virtualized and non-virtualized assets in the datacenter. Working in concert with the Nuage Networks VSP, policies devised for applications automatically extend across virtualized and non-virtualized assets for a fully automated network infrastructure.

FIGURE 2. Nuage Networks datacenter network benefits

	Status Quo	NUAGE NETWORKS DELIVERS What is Needed
Virtualization of network services	LAYER 2 VIRTUALIZATION	FULL NETWORK VIRTUALIZATION, L2 THROUGH L4
Breadth of application models	SIMPLE SCENARIOS	HYBRID CLOUD SERVICES, SEAMLESS VPN CONNECTIVITY
Availability & scale	FRAGILE, NOT MULTI-TENANT	ROBUST, THOUSANDS OF TENANTS
Reach & mobility of network resources	ISLANDS, WITHIN RACKS OR CLUSTERS	SEAMLESS VIRTUALIZED FABRIC, THROUGHOUT & ACROSS DATACENTERS
Network service turn-up time	SLOW, MANUAL, CONFIGURATION DRIVEN	INSTANTANEOUS, AUTOMATED POLICY-DRIVEN
Openness	SPECIFIC TO VENDOR IMPLEMENTATIONS	INDEPENDENCE FROM HARDWARE CHOICES
Breadth of assets automated	VIRTUALIZED ASSETS, LIMITED OPTIONS FOR NON-VIRTUALIZED	ALL DATACENTER ASSETS, VIRTUALIZED & NON-VIRTUALIZED

NU•ÂHJ: FROM FRENCH, MEANING "CLOUD"

The cloud can be more than what it is. In fact, it needs to be. When we founded Nuage Networks, it was with the idea that it's time for the cloud to come of age. From the beginning we recognized the unique challenges that cloud service providers and large enterprises face delivering and managing large, multi-tenant clouds. While the virtualization of compute and storage has evolved quickly, the network simply has not kept up. The result is that today your cloud is being held back. And so is your business.

When we started Nuage Networks, it was with the mission that we could empower our customers to finally deliver on the true promise of the cloud. We envision a world in which IT and IP are no longer in conflict, but rather work in concert to propel your business and elevate the cloud for every one of your customers. We see a world where innovation isn't hampered by infrastructure, and network resources are as effortlessly consumable as compute and storage.

To make this vision a reality, Nuage Networks brings a unique combination of groundbreaking technologies and unmatched networking expertise. This enables us to create solutions that do more than provide incremental improvement. It allows us to introduce radically new thinking and pick up where others have left off, delivering a massively scalable SDN solution that makes the datacenter network able to respond instantly to demand and boundary-less.

Visit us on the web

Our mission is to help you harness the full value of the cloud.

nuagenetworks

www.nuagenetworks.net