Overview: Virtualization takes IT by storm

The adoption of virtualization in datacenters creates the need for a new class of networks designed to support elasticity of resource allocation, increasingly mobile workloads and the shift to production virtual workloads, requiring maximum availability. Building a network that spans both physical servers and virtual machines with consistent capabilities demands a new architectural approach to designing and building the IT infrastructure. Performance, elasticity, and logical addressing structures must be considered as well as the management of the physical and virtual networking infrastructure. Once deployed, a network that is virtualization ready can offer many revolutionary services over a common shared infrastructure.

Virtualization technologies from VMware, Citrix and Microsoft encapsulate existing applications, and abstract them from the physical hardware. Unlike physical machines, virtual machines are represented by a portable software image, which can be instantiated on physical hardware at a moment's notice. With virtualization, comes elasticity where compute capacity can be scaled up or down, on demand, by adjusting the number of virtual machines actively executing on a given physical server. Additionally, virtual machines can be migrated while in service from one physical server to another. Extending this further, virtualization creates 'location freedom' enabling virtual machines to become portable across an ever-increasing geographical distance. As cloud architectures and multi-tenancy capabilities continue to develop and mature, there is an economy of scale that can be realized by aggregating resources across applications, business units, and separate corporations to a common shared, yet segmented, infrastructure.

Elasticity, mobility, automation, and density of virtual machines demands new network architectures focusing on high performance, addressing portability, and the innate understanding of the virtual machine as the new building block of the datacenter. Consistent network-supported and virtualization driven policy and controls are necessary for visibility to virtual machines' state and location as they are created and moved across a virtualized infrastructure.

The Virtual Machine Sprawl Challenge

A direct consequence of the virtual server deployments is the explosion of virtual machines. With new processing architectures it is anticipated to run 10-20 or more virtual machines per server, a significant increase in the number of elements to be managed. This explosion factor will continue to rise, thus further accelerating VM sprawl, as faster and denser multi-core CPUs become the norm in physical servers. Simultaneously, there is a proportional sprawl of virtual switches that VMs connect to within each physical server. Every physical server that hosts VMs has a virtual switch, thus creating a 20-40x increase in the number of network elements to be managed, as shown in Figure 1. Essentially, the network access layer formed by virtual switches is now inside the server.

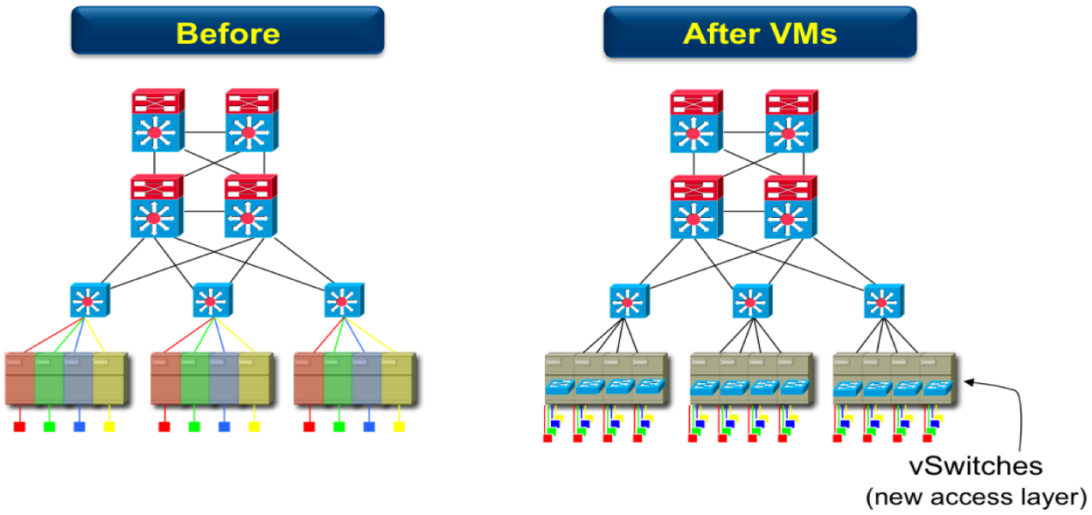

Figure 1: Virtualization creates a new access layer through virtual switches

Virtualization Demands Consistent Cloud Networking

This massive virtual machine and virtual switch infrastructure places new demands on the underlying network fabric for seamless transactions: user-to-VM, VM-to-VM, VM-to-data, VM to Fault Tolerant Peer, and VM Mobility. These new types of transactions demand a network architecture purposefully built to support these demands: a cloud network architecture. Specifically:

1. **VM Explosion:** Since many virtual machines can be instantiated on one physical server, the utilization of physical NIC bandwidth increases proportionally. This NIC link is no longer heavily undersubscribed, which implies that traditional oversubscribed network topologies need to be re-architected for virtualization and private clouds..

   Portable VM images are several gigabytes in size, hence large amounts of data is moved over the network to move and migrate VMs. Also workload elasticity implies that the allocation of virtual machines to compute resource is scaled up or down programmatically, based on various conditions such as load, time of day and power/cooling availability. The network has to be designed with peak performance for this virtualized environment.

2. **Cloud Applications:** New cloud applications integrate Web 2.0 and rich media technologies, often through mash-ups, which can be accessed by millions of users worldwide. This leads to large number of transactions that traverse the network with much higher downstream (VM-to-user) traffic.

   Cloud application workloads are designed to distribute computing tasks across multiple layers

of worker and data nodes, requiring unprecedented VM-to-VM interactions. Also, the RESTful paradigm leads to compute states being kept only on data nodes, thus demanding constant VM access to back-end databases over the network fabric.

3. **VM Migration:** The virtual machine is becoming increasingly mobile, some of the reasons are: retiring/upgrading servers & operating systems, moving workloads off of low-utilization servers so they can be shut off to save power, moving VMs off of a high utilization server for application performance management, or opportunistically migrating workloads to lower-cost compute enclaves. VM mobility requires networks to have larger and flatter Layer-2 domains so the IP address and in-progress client transactions are not disrupted when the VM move executes.

4. **Virtual Switch Management:** Server administrators typically manage virtual networks, because network administrators do not have direct access to built-in virtual switches. For large-scale virtualization and private cloud environments, this creates a major challenge as consistent network-wide policies, monitoring, and diagnostics need to be applied to large number of virtual switches across the infrastructure.

5. **Cloud Reach**: With the emerging offerings from Infrastructure-as-a-Service, public clouds and virtual private clouds, there is currently no way to keep network policy and accounting state intact as the virtual machine moves from one provider to the next. Network state needs to be communicated in a trusted manner to the receiving IaaS cloud provider so the enterprise can retain control over their security policies and enable provider interoperability.

Designing Virtualization-Optimized Cloud Networks

Building the combination of virtual and physical networks to support physical, virtual, and cloud deployments is non trivial. Performance, resiliency, policy control, and management visibility must be considered in the design. The characteristics of networks that support virtualization and cloud computing listed in the earlier section demand that legacy network practices be abandoned in favor of modern cloud networking architectures, such as those enabled by using Arista's 7000 Family of 1Gb/10Gb Ethernet switches and its Extensible Operating System (EOS). Key characteristics of EOS are:

1. **10Gb Ethernet:** High performance 10GbE networking is a must for the core network (with future expansion to 40/100GbE) and in many cases the edge network, especially with blade servers. As VMs drive more bandwidth into the network, millions of users collaborate via centralized cloud applications, and as gigabytes of VM images move across the network fabric, 10Gb Ethernet networking is the only option available today. 10Gb Ethernet is the optimal transport technology to build high performance and highly responsive networks, capable of handling the peak bandwidth demands of cloud and VM workloads. Arista's wire-speed 10Gb Ethernet switches are ideally suited for such performance-demanding environments.

2. **Symmetric cross-sectional bandwidth:** A majority of network architectures today are designed to support one main application: E-Mail. Virtualization and modern application workloads have changed oversubscription rules, which were the basis of legacy network designs. Highly utilized NIC links as well as symmetry in user-to-VM and VM-to-VM traffic require that ingress and egress switching bandwidths be highly balanced; having ingress-to-

egress BW ratio of 1:1 or 2:1 (versus 20:1 or even 40:1 in legacy designs). For instance, the top-of-rack Arista 7048 leaf switch with 48x1GbE server access ports requires 4x10GbE uplinks using 1:1 design rule. For a 48x10GbE switch (whether used as blade server access switch or as a spine switch that aggregates multiple leaf switches) operating at 2:1 oversubscription requires an allocation of 32x10GbE ports for ingress and 16x10GbE ports for egress. The wire-speed architecture of the Arista 7000 Family of 1GbE and 10GbE switches enable deployment of cloud networks with balanced traffic patterns. If you don't oversubscribe, you do not have to manage oversubscription.

3. **Leaf-Spine Architecture:** Constant inter-VM communication and VM mobility drive larger Layer-2 domains, thus driving flatter two-tier leaf-spine architecture over a legacy, enterprise three-tier design. With this architecture, a VM communicates to any other VM in two physical hops or less as shown in Figure 2 below, with Arista's 7000 Family of 1GbE/10GbE Switches.

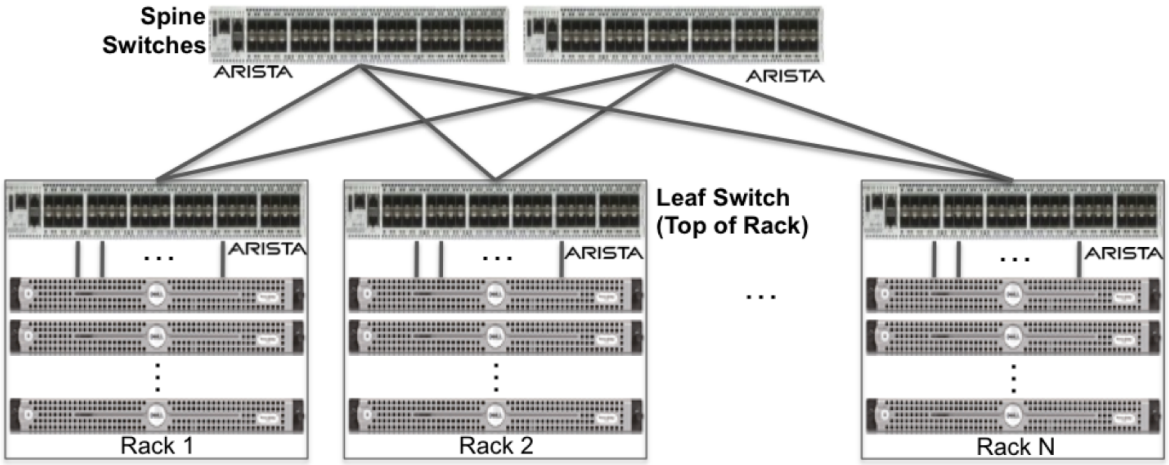

Figure 2: Leaf-Spine Cloud Networking Architecture

4. **Low-latency Switching:** Reducing latency and provisioning proper bandwidth are critical factors for improving application response time. Switches that leverage cut-through packet forwarding modes instead of store-and-forward mode provide 50-90% reductions in per-switch latency. System-wide latency reduction through bandwidth provisioning demands an architectural approach that also includes a 10GbE transport substrate and two-tier network design. Lowering latency improves the efficiency of compute processing and generates business results faster with less power and cost.

5. **Resilient Networking:** Principles of fault-tolerant computing ensure that workloads are not impacted when a few compute nodes (whether physical or virtual) fail. The modern cloud network needs to enable resiliency at the service level. Switches fail mostly because of out-dated operating system and software architectures. Like fault-tolerant compute principles, network operating systems also need to be engineered with fault-tolerant and extensible core operating system design principles. Additionally, a high-speed network control-plane as well as separation of control and data planes has become table stakes for resilient cloud networks.

Arista's EOS is engineered from the ground-up to include these capabilities natively so that the highest levels of network resiliency can be achieved.

6. **Virtualization:** If VMs could easily move from one cloud to another there would be a larger, strategic payoff; yet today's manually managed network is based on static port allocation and static policy definition that has no linkage between the physical and virtual networks and no mechanisms for consistent policy enforcement in the network operating system. For seamless consistency between the virtual and the physical switches, the virtual switches provide transparent redirection using various standards-based mechanisms, including IEEE 802.1Q VLAN tag, MAC address and/or tunnels that are transparent to the physical switches. Proprietary tags need to be avoided as they limit vendor-choice and interoperability.

Consistent network management across both physical and virtual networks demands that heterogeneous virtual switches be managed by network administrators using the well-understood command line interface to simplify adoption yet also provide more programmatic abstractions such as SNMP, XML, XMPP to enable API-based management of the network infrastructure to the Cloud OS. In order to maintain configuration and management consistency across virtual and physical networks as well as during VM migration it is imperative that the management be consistent across physical networks, virtual machines, and cloud implementations.

Introducing Arista's vEOS Solution

Key requirements for enabling seamless operational management across virtual and physical networks are:

- Enable network administrators to treat virtual switches as an extension of the physical network and manage them in a similar manner as physical switches today
- Relieve the burden of network configuration from the VM host administrators
- Provide visibility for troubleshooting and auto-discovery to the network administrator
- Provide a central way of globally implementing and enforcing consistent policy across physical and virtual switches

Arista provides industry-standard interfaces for managing both physical and virtual networks. The Extensible Operating System (EOS) on Arista's 7000 Family of Network Platforms provides management of the physical cloud network. To enable seamless and consistent management of the virtual network, Arista has introduced vEOS, a virtual appliance and a switch, integrated in the implementation of our EOS software. vEOS extends the industry-standard Command Line Interface (CLI) and standards-based management protocols of physical networks to virtualized networks as shown in Figure 3.

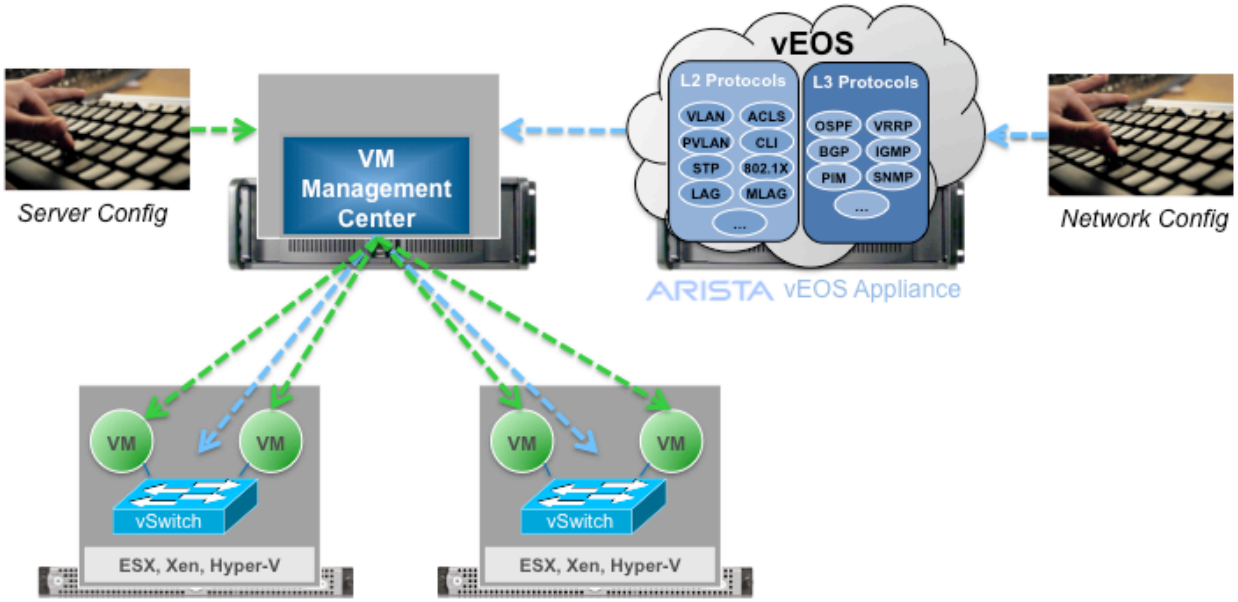

Figure 3: vEOS Management of Physical, Virtual, and Cloud Networks

To a network administrator, vEOS provides the management interface just like a physical switch – SSH/telnet access, industry standard CLI, and integration with SNMP management tools like HP Openview for auto-discovery of VMs. Furthermore, vEOS simplifies network administration tasks such as:

- Configuration of ports groups
- Configuration of uplinks
- Monitoring virtual switches, operation and status
- Creation of distributed port profiles for consistent management across physical and virtual networks

With vEOS, both server and network administrators can manage the virtualized cloud infrastructure in tandem; using their own unique and well understood workflows.

Another key benefit of vEOS is that it automates network provisioning during VM migration: P2V, V2V and V2C. When a VM migrates from physical port to virtual or from one virtualized host to another across the network, vEOS migrates the corresponding network settings providing end-to-end consistency. While other solutions require replacing the existing internal virtual switch with vendor-specific product, causing outage during the insertion, Arista's EOS is unique in its ability to work with existing virtual switches while providing a consistent administrative experience.

Summary

The combination of virtualization and cloud computing creates a computing paradigm that requires careful network considerations. A cloud network must support the abstraction of virtualization services from specific physical servers. In particular, 10GbE networking and orchestration mechanisms must allow on-demand deployment of network bandwidth & virtual machine resources and support isolation between different workloads and customers.

The migration from network silos to virtualization and clouds signifies the replacement of vastly oversubscribed switches, with network infrastructure designed for virtualization and the cloud such as Arista's 7000 Family deployed in two-tiered leaf and spine designs. The Arista 7000 Family along with EOS and vEOS plans to support smooth migration of VMware's vSphere/vNetwork Distributed Switches, Citrix's Xen, Kernel Virtual Machine, and Microsoft's Hyper-V hypervisors. Orchestration is achieved via industry standard discovery protocols, CLIs and extensible APIs. Undoubtedly the physical, virtual, and cloud networks must not only co-exist but also be managed seamlessly for scalable deployments.

The combination of vEOS and the Arista 7000 Family enables physical network and virtualization operators as well as private and public cloud systems to seamlessly migrate virtual machines across network, server, and organizational boundaries with consistent policy and accounting. It is a dramatic departure from the silos of today as workloads migrate from physical to virtual to cloud networking.