Storage. Networking. Accelerated.™

Whitepaper

Virtualization of Datacenters: Network Challenges and Control Plane Solutions

Raghu Kondapalli, Director of Technology

May 2012

This paper discusses the network-related challenges in virtualized datacenters. It focuses on their control plane requirements, and outlines effective solutions to satisfy to these requirements and improve datacenter network efficiency.

Introduction

Today's content-rich Internet, e-commerce, and social media applications require instantaneous and reliable availability to the massive amounts of data served on a daily basis. This has generated an unprecedented demand for the processing, storing, and transporting of terabytes of data. Ongoing migration to cloud computing in the business, finance, and consumer world indicates the universal adoption of the datacenter-based business model in the IT industry. Datacenters need to be cost-effective, power-efficient, scalable, and agile to satisfy the 24/7 data access demand. This is the impetus for the datacenter transition from the physical to the virtual world. Datacenter virtualization started with server virtualization - technologies such as multicore CPU systems and multi-thread operating systems enabled initial server virtualization. The need for storing and managing massive amounts of data in an efficient manner mandated the virtualization of storage area networks (SAN) and network attached storage (NAS). However, because of the need for a maximum return on investment (ROI) on datacenters, efficient transport of data is critical — which is driving the virtualization of the datacenter networks.

This paper outlines the demand creators and current solutions for datacenter network virtualization. It discusses the key technical challenges that need to be addressed to transform the datacenter networks for maximum utilization, and how control plane scaling is a key component in the complete virtualization of datacenter networks. This paper describes the network-related challenges faced by the existing virtualized datacenters and highlights the control plane architectures needed to address them.

Datacenter Virtualization - Demand Generators and Solutions

The need for instantaneous and reliable availability of data across all segments of today's connected world – from e-commerce, business, medical, finance, education, and communication to social media, entertainment, and gaming – is pushing the boundaries of datacenter virtualization The cloud computing phenomenon, with its scalability, differentiation, and lower total-cost-of-ownership (TCO), requires virtualization in all components of the datacenter. Cloud-based enterprise web services, such as Elastic Compute Cloud (EC2) from Amazon, Microsoft®'s online services, and consumer applications such as iCloud from Apple and Web2.0, are catalysts in the adoption of virtualization, specifically multi-tenancy, by datacenter operators.

Servers within datacenter were virtualized by implementing virtual machines (VMs). On the software side, hypervisors created abstraction between physical and virtual machines and they absorbed many of the connectivity, manageability, and scalability aspects. The open source hypervisor Xen and products such as Microsoft® Hyper-V Server enabled the distribution of workloads from physical server to several VMs. Hypervisor-based server platforms became the dominant solution for datacenter virtualization and scalability. However, software-based hypervisors are unable to keep up with the performance demands of datacenters' growing capacities. In mid 2000s, processor industry leaders deployed hardware extensions to support x86 processor virtualization. One such example is the AMD-V technology. These technologies provided the hardware acceleration needed to support virtualization and enabled multiple operating systems to share the processor resources.

Server virtualization affected datacenter storage – these compute VMs typically reside on a SAN. They initialize after VM images are loaded into server memory, which then means the SAN needs extra storage to replicate and backup the VM images dynamically. Initial phases of storage transformation resulted in SANs and storage hypervisors. This helped the storage administrator perform the tasks of backup, archiving, and recovery more easily, and in less time, by disguising the actual complexity of the SAN. However, these techniques by themselves were not able to keep up with the datacenter storage demands – hardware acceleration was critical to boost SAN performance and utilize storage virtualization. Storage solutions, such as the LSI® Nytro™ application acceleration solutions, successfully addressed this challenge. These application-aware storage solutions provide

(cont. p.2)

hardware acceleration to SAN and directly attached storage (DAS), and enabled data-center SAN infrastructure to speed-up database accesses and provide real-time analytics.

Until recently, most of the efforts in datacenter virtualization addressed the server and storage segments. Network virtualization was adhoc and implemented as an add-on module to traditional compute-centric hypervisors. Network-specific extensions to hypervisors managed the basic connectivity and fault management, and were able to meet the performance needs for small datacenters with limited VMs per server platform. However, the current generation of server farms implements thousands of servers and many VMs per server. The application workloads, which are distributed across several VMs, increase VM-to-VM communication (east-west traffic). Other factors, such as VM migration and storage applications such as data replication, have also increased east-west traffic. Now, the datacenter network optimized for server-to-client traffic (north-south traffic) needs to be transformed in order to reap the full benefits from virtualizing the datacenter.

At present, several solutions are proposed to improve datacenter network utilization. At the network architecture level, isolating the control plane functions from the data plane and virtualizing them is a growing trend. It involves improving the efficiency of the existing network infrastructures with simple upgrades. Scale-out and scale-up are two such techniques that are being considered in the industry. In the scale-out approach, the control plane functions are separated and distributed across a server network. Figure 1 illustrates the scale-out approach. Here the control plane runs on a separate server located in the server farm or virtualized to reside in a cloud. Google is an example of a mega datacenter that has adopted the scale-out model.

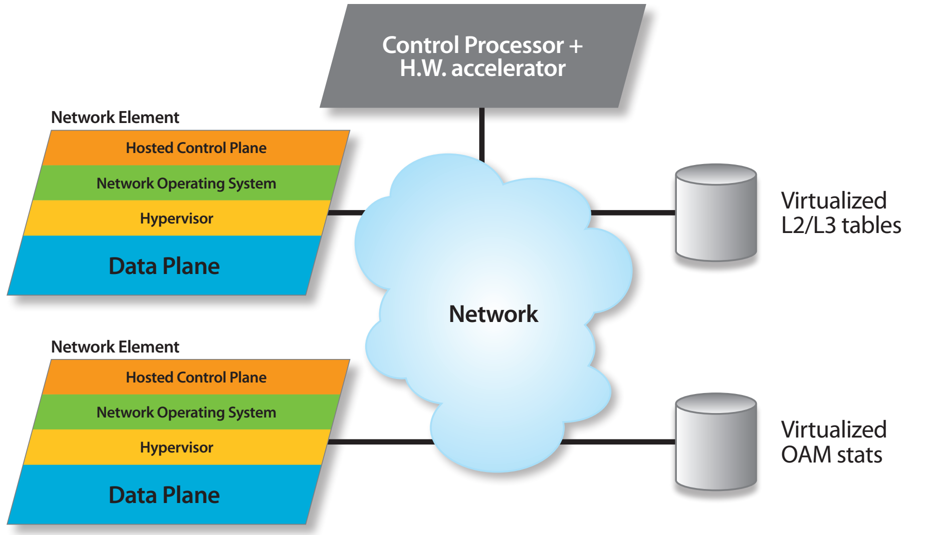

Figure 1: Control plane scale-out architecture

In the scale-up approach, the server's processing power is augmented by adding extra compute resources, such as x86 processors, to existing networking systems (as shown in Figure 2). This approach is often implemented by traditional datacenter networks such as financial institutions. In both the scale-out and scale-up architectures, performance is enhanced further by providing function-specific hardware acceleration.

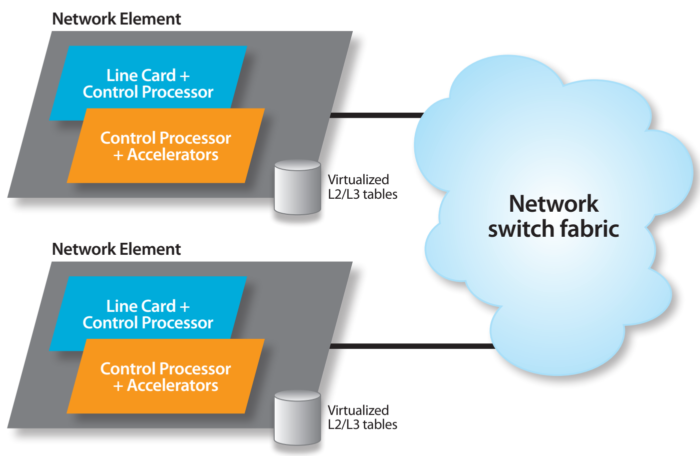

Figure 2: Control plane scale-up architecture

Another trend, primarily driven by academic research, is Software-Defined Networking (SDN). It is a new network design concept that decouples the software that directs traffic and physical network infrastructure. Using SDN abstraction, the network application stacks are presented with a virtual view of the network instead of the physical topology. This enables the faster completion of network events such as VM migrations in virtualized networks. SDN allows virtualizing and distributing the control plane tasks across a network. SDN is implemented using a protocol such as OpenFlow. OpenFlow is a programmable network protocol, being developed and standardized by the Open Networking Foundation (ONF), to manage traffic among switches from various vendors. It proposes to separate control plane functions, such as routing, from data plane functions, like forwarding, and allow them to execute independently on different devices.

Network Challenges in Virtualized Datacenter

This section outlines the network management complexities in a virtualized datacenter and highlights the control plane requirements needed to address these challenges.

Server virtualization overheads

Server virtualization has enabled hundreds of VMs per blade server in a datacenter using multicore CPU technology. As a result, packet processing functions, such as packet classification, routing decisions, encryption/decryption, and security decisions, have increased exponentially. Discrete devices cannot be scaled to meet these demands and they need to collapse to virtual platforms. These packet-processing functions, implemented in software as network hypervisors, are not efficient since the x86 servers are not optimized for these network-processing functions. The control plane needs to be scaled by adding processors dedicated to perform network control tasks. It also needs hardware assistance provided by the function-specific control plane accelerators.

Traditional Datacenter	Virtualized Datacenter
10K servers	Assuming 4 VMs per core and 32 cores per server → 128 VMs per server x 10 K servers → 1.28M VMs or distinct addressable virtual servers
10 K NICs	1.28 M virtual NICs & 10 K physical NICs
Using 48-GE switches → 208.3 switches	Using 48-GE switches → 26,666 virtual switches and 208.3 physical switches
Assuming network device & element management is 1% traffic overhead → 480 Gbps x 208 x 0.01 = 998 Gbps → 1 Tbps mgmt. overheads (approx.)	Assuming only 25% VM-to-VM switching and control; network device & element management is 1% traffic overhead → 480 Gbps x (26,666*0.25 + 208) x 0.01 = 32 Tbps → 32 Tbps mgmt. overheads (approx.)

Table 1: Comparison of virtualized datacenter network management overheads with a physical datacenter

Table 1 compares packet-processing overheads between a traditional, physical datacenter and datacenter with server virtualization. By mapping one physical server to four VMs and assuming 1% traffic management overheads with 25% east-west traffic, the network management overhead increases 32-fold.

Virtual machine (VM) migration

Support for VM migration between servers, either within one server cluster or across multiple clusters, is another key challenge faced by network managers. An IT administrator may decide to move a VM from one server to another for a variety of reasons including resource availability, quality-of-experience, maintenance, and hardware/software or network failures. The hypervisor handles these VM migration scenarios by first reserving a VM on the destination server, moving the VM to the destination, and finally tearing down the source VM. Hypervisors are not capable of the timely generation of address resolution protocol (ARP) broadcasts to notify the VM moves, especially in huge VM networks. Situations like congestion (in which ARP messages cannot get through the network in a timely manner) and network re-configuration during a VM migration can cause huge control overheads. The control plane needs to manage the dynamically changing network behavior – the specific requirements include fast ARP message generations, connection tear-up, and routing decision updates.

Multi-tenancy & security

Due to the high cost associated with building and operating a datacenter, IT organizations are moving towards the multi-tenant datacenter model – where different companies or departments share the same infrastructure and virtual resources. Data protection and isolation is a critical requirement of a multi-tenant datacenter. Multi-tenancy requires logical isolation of resources without dedicating physical resources to each customer, and the network control needs to implement customer-specific policies and Quality-of-Service (QoS) definitions. The control plane needs to provide secure access to datacenter resources and dynamically change the security policies based on VM migration.

Service-level agreements (SLA) and resource metering

Resource metering by collection of network statistics is essential for IT organization – it is used for datacenter ROI computation, infrastructure upgrades, and expansion. At present, the network monitoring tasks are spread across the hypervisor, legacy management tools, and newer infrastructure monitoring tools. It is critical that that the control plane collects and consolidates the management information.

Control Plane Solutions for Virtualized Datacenters

The key trends in control plane re-architecture were introduced earlier, now we will discuss the solutions to satisfy these requirements.

Control plane scale-up solution

In this method, the existing network control platforms are supplemented by additional or more powerful compute engines, to help execute networking control stack. An example configuration was shown in Figure 2. This method frees up the server CPU cycles resulting in the overall improvement of the network performance. However, since general-purpose processors are not optimized for the network tasks, they are not an ideal solution. Under heavy or bursty traffic conditions, the network will experience significant performance degradation. This deficiency is solved by the addition of a function-specific, protocol-aware control plane processor to the control card.

(cont. p.4)

The solution can be implemented in a legacy infrastructure as an add-on acceleration card or as an accelerator System-on-a-Chip (SoC) for new platforms. Another control plane scale-up technique is to add the control coprocessor to offload the routine control plane tasks to the data line card itself. In this case, the central control card is used for more specialized tasks such as routing decisions. Utilizing the central control plane card with accelerated SoCs can reduce the turnaround latencies and also increase the system reliability.

Control plane scale-out solution

In the scale-out architecture, the basic platform is implemented with generic processors and augmented with separate function specific engines like a network control plane processor. The control plane tasks are broken into sub-components, such as discovery, dissemination, and recovery, and these components are virtualized across the datacenter, as described in the SDN approach. Figure 1 illustrates the scale-out approach. Such a control plane can execute on any server in the network or in the cloud. It communicates with data planes using APIs based on the network protocol, such as OpenFlow. Based on the network traffic needs, specific control plane tasks are accelerated. For example, if there are IT scheduled network moves, the discovery and dissemination components are scaled to meet the processing needs.

The advantages include ease of scaling the control engine and isolation of network management without incurring control plane overheads. It relies on robust communication between the control plane and the data planes. In addition, hardware acceleration is necessary to boost the control plane server performance.

Function-specific communications processors are designed to handle specific control plane tasks such as ARP offload, OAM offload, security, network statistics, QoS, and providing hardware acceleration. The LSI Axxia® Communication Processor family is an example of protocol-aware communications processor. It provides acceleration for a variety of network management functions including packet analysis and routing, security, control plane ARP, IGMP messages, networking statistics, application aware firewall, and QoS. Its flexible data plane processing enables support for new applications such as OpenFlow and SDN. Its flexible data plane processing enables support for new applications such as OpenFlow and SDN. Hardware acceleration is essential in control plane scale-up and scale-out architectures implementations for optimum network utilization.

Conclusion

Datacenters are going through transformations to meet the exploding data processing, storage, and transport needs. Server and storage virtualization, cloud-based service application models, and operational costs have resulted in increased VM server densities, VM-to-VM communication, VM migration, and multi-tenancy. Server and storage virtualization technologies, implemented as a hybrid of software-based hypervisors and hardware accelerators, have significantly increased datacenter utilization. However, to achieve a maximum ROI from a datacenter, the current datacenter network infrastructure has to transform. Network virtualization techniques based on hypervisors are not an optimum solution to meet the growing network traffic. The control plane architecture is being redefined to solve these current network inadequacies. The solutions include scale-up and scale-out techniques – scale-up being based on adding more resources to existing network infrastructure; scale-out being based on virtualizing and farming out control plane functions to virtualized compute engines. Newer network design concepts, such as SDN along with the OpenFlow protocol, support these control plane virtualization techniques. The control plane scaling technologies, enabled by the state-of-the-art control plane communication processor SoCs such as the Axxia Communication Processor provide efficient solutions for maximum utilization of a datacenter network.

Storage. Networking. Accelerated.™

For more information and sales office locations, please visit the LSI web sites at: lsi.com lsi.com/contacts

North American Headquarters
Milpitas, CA
T: +1.866.574.5741 (within U.S.)
T: +1.408.954.3108 (outside U.S.)

LSI Europe Ltd.
European Headquarters
United Kingdom
T: [+44] 1344.413200

LSI KK Headquarters
Tokyo, Japan
Tel: [+81] 3.5463.7165