IMPLEMENTING VMWARE SERVER VIRTUALIZATION ON JUNIPER NETWORKS INFRASTRUCTURE

Although Juniper Networks has attempted to provide accurate information in this guide, Juniper Networks does not warrant or guarantee the accuracy of the information provided herein. Third party product descriptions and related technical details provided in this document are for information purposes only and such products are not supported by Juniper Networks. All information provided in this guide is provided “as is”, with all faults, and without warranty of any kind, either expressed or implied or statutory. Juniper Networks and its suppliers hereby disclaim all warranties related to this guide and the information contained herein, whether expressed or implied of statutory including, without limitation, those of merchantability, fitness for a particular purpose and noninfringement, or arising from a course of dealing, usage, or trade practice.

Copyright © 2009, Juniper Networks, Inc.
# Table of Contents

- Introduction .................................................................................................................. 4
- Scope ............................................................................................................................... 4
- Target Audience ............................................................................................................ 4
- Terminology and Concepts ............................................................................................. 5
  - Server Virtualization .................................................................................................... 5
  - Network Virtualization ............................................................................................... 5
- Design Considerations .................................................................................................... 7
  - Juniper Networks Data Center Network Reference Architecture .............................. 7
- Switching Design—Top-of-Rack ..................................................................................... 9
- Top of Rack—EX4200 Switches ..................................................................................... 9
- EX4200 Virtual Chassis Technology and Its Application for Virtualized Servers .......... 9
- High Availability at the Core and Access Layers .......................................................... 10
- VMotion Traffic Flow Considerations ........................................................................... 11
  - VMotion Traffic Across the Access and Core Layers ................................................ 12
  - VMotion Traffic in the Access Layer Within the Same Rack (128 Gbps Virtual Backplane) ................................................................. 12
  - VMotion Traffic in the Access Layer Across Multiple Racks (10GbE Uplink) ........ 13
- How VMotion Works and Best Practices ..................................................................... 13
- Implementation ................................................................................................................ 15
  - Configuring VMware Connectivity to the Access Layer Interconnecting to the Core Layer ........................................................................................................................ 15
  - Data Center Access Layer and Core Layer Connection Recommendations ................ 16
  - VMware Server Connection Recommendations ....................................................... 17
  - Configuring EX4200 Access Switches ....................................................................... 17
  - Configuring the EX8200 Ethernet Switches ............................................................... 18
  - Configuring the MX960 Ethernet Services Router .................................................... 18
  - Configuring the Virtual Switch on VMware ESX ....................................................... 18
  - General Recommended Steps for Configuring a Virtual Switch on VMware ESX ....... 19
  - Configuring VMware Connectivity at the Access Layer Within the Same Rack (128 Gbps Virtual Chassis Backplane) ................................................................. 19
  - General Recommended Steps for Configuring the EX4200 Access Port for VMotion ......................................................................................................................... 20
  - Configuring VMware Connectivity at the Access Layer Across Multiple Racks (10GbE Uplink) ................................................................. 20
  - General Recommended Steps for Configuring Uplink for Virtual Chassis Technology ......................................................................................................................... 21
- Summary ........................................................................................................................... 22
- Appendix A: Code Samples ............................................................................................. 23
  - EX4200 Spanning Tree Configuration ......................................................................... 23
  - EX8200 Configuration ................................................................................................. 23
  - MX960 Configuration .................................................................................................... 24
- Appendix B: VMware Key Concepts and Terms .............................................................. 26
  - Table 1: Key Terms and Definitions for VMware’s ESX 3.5 Platform ....................... 26
  - About Juniper Networks .............................................................................................. 28
Table of Figures

Figure 1: EX4200 and EX8200 switches and MX960 router positioned in access and core layers ........................................... 4
Figure 2: VMware ESX 3.5 server, VMware virtualization layers ................................................................................................. 5
Figure 3: VLAN tagging—VGT, EST and VST modes .................................................................................................................. 6
Figure 4: Juniper Networks data center reference architecture .................................................................................................. 7
Figure 5: Connectivity between server racks and EX4200 switches .......................................................................................... 8
Figure 6: EX4200 switch in the application and data services network ....................................................................................... 8
Figure 7: Scaling through EX4200 Virtual Chassis switch top-of-rack configuration ................................................................. 9
Figure 8: EX4200 switches implemented as top-of-rack switches with VMware connection .................................................... 10
Figure 9: Example of high availability and load balancing in core and access layers ................................................................. 11
Figure 10: VMotion traffic flows across the access layer (Virtual Chassis) and core layer .......................................................... 12
Figure 11: VMotion traffic flows in the access layer within the same rack (128 Gbps Virtual Chassis backplane) ......................... 13
Figure 12: VMotion traffic flow across GbE uplinks to different racks ....................................................................................... 13
Figure 13: VirtualCenter/production network ........................................................................................................................ 14
Figure 14: Interconnecting VMware ESX server, EX4200 switch, EX8200 switch or MX960 Ethernet Services Router ................. 15
Figure 15: Connecting VMware ESX server to access and core layer ......................................................................................... 16
Figure 16: Example of VMware ESX server NIC teaming to the access layer ........................................................................... 17
Figure 17: VMware ESX virtual switch configuration ................................................................................................................ 19
Figure 18: VMware connectivity across 128 Gbps Virtual Chassis backplane ..................................................................... 20
Figure 19: VMware connectivity across multiple racks (using 10GbE uplink) ........................................................................ 21
Figure 20: Virtual Chassis racks across 10GbE .......................................................................................................................... 21
Introduction

In today’s enterprise data centers, servers, applications and networking platforms are being consolidated. Network technology is being standardized to reduce costs, increase efficiencies and improve resource utilization. Virtualization is a key technology that meets these requirements. Applications are no longer bound to specific physical resources, but instead rely on a pool of CPUs, memory and network infrastructure services made available through virtualization.

This paper provides various server and network virtualization design considerations typically associated with the data center environment. These design considerations are based on deploying Juniper Networks® EX4200 line of Ethernet Switches with Virtual Chassis technology, EX8200 modular chassis switches and Juniper Networks MX Series Ethernet Services Routers. We provide specific design considerations and implementation guidelines around VMware implementation in the data center network. We address specific high availability (HA) scenarios with special attention given to the flexibility of the Virtual Chassis technology-based architecture. Design and optimization steps address VMware’s ESX 3.5 virtual server network deployment, scalability and redundancy advantages.

The implementation guidelines cover the following major topics:

- Configuring EX4200 and EX8200 switches and MX Series routers to integrate with VMware’s ESX virtual server for establishing a flat Layer 2 domain across the aggregated server pool
- Configuring the EX4200 Virtual Chassis technology to implement in the access layer
- Configuring a 10GbE uplink from an EX4200 switch

Figure 1 shows an example of the EX4200 and EX8200 switches and the MX960 Ethernet Services Router in the access, aggregation and core layers.

Figure 1: EX4200 and EX8200 switches or MX960 router positioned in access and core layers

Scope

This paper presents design considerations, best practices and implementation guidelines for integrating VMware’s ESX 3.5 infrastructure into the enterprise network, in conjunction with deploying EX4200 switches at the access layer and EX8200 switches or MX Series routers at the core layer based on Juniper Networks data center network reference architecture.

Target Audience

- Security and IT engineers
- Network architects
Terminology and Concepts

The following terms represent key items and concepts related to virtualization. These will serve as an aid in helping you understand virtualization as it applies to the design considerations and implementation guidelines presented in this paper. For further details concerning VMware-related concepts and additional documentation, see Appendix B, VMware Key Concepts and Terms.

Server Virtualization

Server virtualization is the masking of server resources, including the number and identity of individual physical servers, processors and operating systems, from server users. The server administrator uses a software application to divide one physical server into multiple isolated virtual environments. The virtual environments are sometimes called virtual private servers, but they are also known as partitions, guests, instances, containers or emulations. There are three popular approaches to server virtualization: the virtual machine model, the para-virtual machine model, and virtualization at the operating system (OS) layer, as shown in Figure 2.

The virtual machine model (addressed in this paper) is based on the host/guest paradigm. This model permits dissimilar operating systems to run side-by-side on the same hardware by running under a virtual machine, which effectively tricks OS running under it into thinking it is running by itself on the entire native hardware. This approach allows the guest operating system to run without modifications. VMware and Microsoft Virtual Server both use the virtual machine model. To stay within the scope of this paper, we only address the concept of server virtualization in the virtual machine model.

![Figure 2: VMware ESX 3.5 server, VMware virtualization layers](image)

Network Virtualization

Network virtualization is the process of abstracting a single physical network, so as to permit its shared use by several logical entities—thus ensuring efficient utilization of that physical network. Network virtualization allows enterprises to segment the network using “virtual” objects, such as virtual LANs (VLANs), virtual routers, security zones, virtual route forwarding and virtual systems. These features can help enterprises take advantage of the efficiencies and security of network segmentation, without significantly increasing cost or network complexity and administrative burden.

VLANs provide for logical groupings of stations or switch ports, allowing communications as if all stations or ports were on the same physical LAN segment. Confining broadcast traffic to a subset of switch ports or end users saves significant amounts of network bandwidth and processor time.

VLANs are used in widely distributed computing environments as a means to identify and then segment traffic at a very granular level. Within Juniper Networks appliances, VLANs are used as a complement to virtual systems and security zones, segmenting traffic with a VLAN tag, and then routing that traffic to a security zone, a virtual system or a physical interface based on a security policy.
Security Zones allow enterprises to divide the physical network into a series of virtual sections, so that they can establish various levels of trust and contain attacks. With Juniper Networks security zones, enterprises can define areas as specific as the finance department and wireless LAN users or as general as the untrust, trust and demilitarized zones (DMZ). Unlike some competitive security solutions that rigidly support untrust, trust and DMZs, each of the security zones in Juniper’s products can be customized to suit specific security needs.

Multiprotocol Label Switching (MPLS) is a standards-approved technology for speeding up network traffic flow and making it easier to manage. MPLS involves setting up a specific path for a given sequence of packets, identified by a label put in each packet, thus saving the time needed for a router to look up the address of the next node in order to forward the packet.

Label switched path (LSP) is a sequence of hops (R0...Rn) in which a packet travels from R0 to Rn through label switching mechanisms. A label switched path can be chosen dynamically, based on normal routing mechanisms or through configuration.

In the enterprise core network, Juniper Networks recommends a two-layer aggregation network model as opposed to the conventional three-layer aggregation network model. By combining the EX4200 in the access layer and line-rate EX8200 switches or MX Series Ethernet Services Routers in the aggregation/core layers, we can collapse the switching tiers in the data center LAN. The EX4200 switch allows us to combine multiple racks and multiple rows into a single logical switch for administrative purposes. Uplinks from several racks now become optional, and this reduces the number of connection points required in the aggregation tier. The EX8200 switches and MX Series routers allow us to deliver line-rate 10GbE at scale and full performance with all features turned on. This by itself requires fewer ports, and hence switches, compared to other switches in the marketplace. The overall reduction in the number of connection points in the aggregation tier allows the Juniper Networks architecture to reduce the total number of switches in the aggregation tier and, in several cases, reduce the number of tiers in data center LAN switching.

From a network virtualization perspective, there are various technologies that provide data, control and management plane virtualization. A data plane virtualization example is a single physical interface to provide security to multiple network segments using 802.1q VLAN tagging. From a control plane virtualization perspective, multiple routing domains and protocol instances are other examples. A management plane virtualization example supports multiple logical firewall/VPN security systems that use virtual systems (VSYS) for true multi-department or multi-customer environments, such as large enterprises or service providers who offer managed security services all in a single device. Figure 3 illustrates VLAN tagging and emphasizes its three modes—Virtual Switch (internal to the hypervisor), EX4200 switch and Virtual Guest tagging.

![Figure 3: VLAN tagging—VGT, EST and VST modes](image)

**NOTE:** The dashed lines indicate the origin and direction of VLAN tagging.

For additional VMware concepts and terms, see *Appendix B VMware Key Concepts and Terms* for definition of major terms used in VMware nomenclature discussed in this paper. Also visit [www.vmware.com/products/vi/vc/vmotion.html](http://www.vmware.com/products/vi/vc/vmotion.html).

We will now show you how Juniper’s networking devices are leveraged by Juniper Networks reference architecture in the following design considerations section.
Design Considerations

Now let’s talk about how virtual networking and all of its major components—the virtual switch (internal to the hypervisor), NIC teaming, and VMotion apply to the Juniper Networks data center reference architecture. For more information about VMware ESX 3 server virtualization, see Appendix B VMware Key Concepts and Terms.

Juniper Networks Data Center Network Reference Architecture

Figure 4 depicts a high-level view of Juniper Networks data center reference architecture. The core/aggregation network tier connects to the Applications and Data Services tier which hosts all of the servers, databases and storage. The Edge Services tier provides connectivity to the Internet and WAN as well as edge security services, including VPN termination.

Figure 4: Juniper Networks data center reference architecture
In this paper, we primarily discuss the connectivity options for the Applications and Data Services tier, as shown in Figure 4. In data center environments, servers are interconnected to access switches deployed within server racks. These access switches are often referred to as “top-of-rack” switches due to their location within the data center. Top-of-rack switching, implemented with an EX4200 switch, provides increased levels of availability because of multiple independent operating characteristics and physical power sources. Servers connect to two different physical switches, each part of a separate Virtual Chassis ring. Each ring in turn connects to the core network while using a loop detection and HA Layer 2 protocol. See Figure 5.

![Figure 5: Connectivity between server racks and EX4200 switches](image)

As shown in Figure 6, each internal and external applications network can be segmented into several sub networks. The servers that host these applications connect with 1 GbE links to the EX4200 switch with Virtual Chassis technology. The EX4200 switch connects to the network core/aggregation layer through a 10 Gbps connection. Depending on the number of servers, multiple EX4200 switches might be required, as shown in Figure 6. Juniper Networks recommends dual homing the access layer switches (EX4200) using Layer 3 with OSPF equal-cost multipath (ECMP) instead of using the Spanning Tree Protocol (STP) for deterministic behavior for minimal packet loss.

**NOTE:** The above mentioned design option using Layer 3 with OSPF ECMP on the access layer switch is not covered in this document. This paper is structured to present VMware use cases and deploy Layer 2 on the access layer EX4200 switches.

![Figure 6: EX4200 switch in the application and data services network](image)

**NOTE:** This recommendation applies except when the aggregated server pool over which you are running live migration does not lie entirely within a Virtual Chassis and you need to extend the Layer 2 domain to the EX8200 switches or MX960 routers in collapsed aggregation and core layers.
**Switching Design—Top-of-Rack**

The “top-of-rack” EX4200 switch with Virtual Chassis technology configuration is the primary design consideration in a Virtual Chassis design for connecting virtualized servers. This section defines and discusses how this configuration applies to VMware virtualization and the EX4200 switch.

**Top of Rack—EX4200 Switches**

The EX4200 switch eliminates the cable management headaches posed by “end-of-row” chassis-based platforms. It also eliminates the need for multiple, commonly used “top-of-rack” stackable switches—an architecture traditionally used to provide a high availability design.

In the following “top-of-rack” configuration, a pair of stackable EX4200 switches with Virtual Chassis technology typically sits atop a rack of servers (multiple servers can also be stacked in a “top-of-rack” formation). For redundancy purposes, each VMware ESX physical server in the rack supports two connections, one to each Virtual Chassis, as shown in Figure 7. In this configuration, EX4200 switches with Virtual Chassis technology are stacked on top of each other and are connected to the core/aggregation switches in a dual-homed manner, providing the necessary “always-on” link between users and critical backend VMware ESX physical servers.

![Figure 7: Scaling through EX4200 Virtual Chassis switch top-of-rack configuration](image)

The Juniper Networks Virtual Chassis technology enables a data center to scale and add/stack as many EX4200 switches (up to 10 in a Virtual Chassis configuration) as needed to meet connectivity needs, while delivering true chassis-like functionality.

For a more resilient design expanding port density, up to 10 EX4200 switches can be interconnected over a 128 Gbps backplane, thereby creating a single logical switch that supports up to 480 10/100/1000BASE-T ports and up to 40 GbE or 20 10GbE uplink ports using the uplink module.

For more information about the benefits of top-of-rack and end-of-row deployment, refer to the Juniper Networks Data Center LAN Connectivity Design Guide.

**EX4200 Virtual Chassis Technology and Its Application for Virtualized Servers**

With the EX4200 switches, IT gets the best of both worlds—deployment flexibility and the economics of stackable platforms combined with the advanced Layer 2 though Layer 4 functionality and the carrier-class reliability of a high-end chassis platform. While the redundant top-of-rack configuration, as shown in Figure 8, provides resiliency against equipment failures, traditional stackable switches without Virtual Chassis technology in general lack many of the HA characteristics associated with chassis-based switches, such as redundant power supplies, redundant cooling fans and fast, seamless failover.
High Availability at the Core and Access Layers

The core network is a key component in enabling HA in the data center network. By connecting all networks to the core network with full redundancy at the core, HA is achieved without added complexity and dependency on network protocols and convergence. Traditionally, adding HA required redesign of the network, whereas by using a core network approach and standards-based redundancy protocols such as VRRP (Virtual Router Redundancy Protocol), OSPF or other open-standard protocols for rapid such-second convergence, HA is provided at easier operational overhead. In addition to adding redundant devices, it is extremely important to ensure that the core data center devices support in-service operations such as hot-swap interfaces and software upgrades. In Juniper Networks data center network reference architecture, the core network is implemented with either the EX8200 switches or MX Series routers that are capable of providing a very high level of performance, HA and network virtualization features.

At the access layer, EX4200 switches offer redundant hot-swappable power supplies and a field replaceable fan tray. In addition, when two or more EX4200 switches are interconnected to form a Virtual Chassis configuration, Juniper Networks Junos® Software makes use of the multiple Routing Engines to deliver graceful Routing Engine switchover (GRES), as well as Layer 2 non-stop forwarding in the event of a Routing Engine failure. This ensures that network operations continue uninterrupted and that no critical routing data is lost following a primary Routing Engine failure. Primary and backup Routing Engines are automatically assigned by Junos, dictating an orderly transfer of control-plane functions.

Figure 9 illustrates an example of the EX4200 access switch that provides HA and load balancing at the access layer (Virtual Chassis) along with the EX8200 Ethernet switches or MX960 Ethernet Services Router, which lies in the core layer.
Figure 9: Example of high availability and load balancing in core and access layers

As illustrated in Figure 9, two redundant VMware ESX servers are placed in two separate racks and are interconnected to the data center infrastructure access layer, using two EX4200 switches with Virtual Chassis technology on each rack as “top of rack” switches, where each EX4200 switch connects with an uplink to the data center’s core/aggregation layer that runs the EX8200 switch or MX960 router.

The key reason for this deployment is a redundant configuration (each Virtual Chassis, Rack 1 and Rack 2, backs up the other). VMware's default mode of operation utilizes both links at any time, distributing the traffic across links based on network interface card (NIC) teaming and port group configuration.

NIC teaming is a VMware feature where you can connect a single virtual switch (internal to hypervisor) to multiple physical Ethernet adapters. See Appendix B VMware Key Concepts and Terms for a detailed definition on NIC teaming. A team can share the load of traffic between physical and virtual networks among some or all of its members, and it can provide passive failover in the event of a hardware failure or a network outage. To configure for NIC teaming, see Configuring the EX4200 Access Switches.

Each VMware ESX server physical interface connects to a different EX4200 Virtual Chassis configuration. See Appendix A Code Samples which provides the configuration code samples of the VMware networking relevant to this configuration and shows the load balancing characteristics that are associated with different configurations.

Each EX4200 Virtual Chassis configuration connects through an uplink to either the EX8200 or MX960 core tier. The core tier EX8200 or MX960 links the Virtual Server Networks, the VirtualCenter Network, and the VMotion Networks across all Virtual Chassis. The EX4200 switch configuration matching with the VMware port group setup is also discussed in the following configurations.

VMotion Traffic Flow Considerations

This section primarily covers the following three scenarios for design considerations that focus on establishing VMotion traffic flow:

- VMotion traffic across the access and core layers
- VMotion traffic in the access layer within the same rack (128 Gbps virtual backplane)
- VMotion traffic in the access layer across multiple racks (10 Gbps uplink)
On VMware enterprise servers such as the ESX 3.5 server, VMotion is a feature that supports the live migration of running virtual machines between physical servers. It is a solution to eliminate planned downtime. Most planned downtime is for hardware maintenance activities such as memory and storage upgrades, power supply replacements or fan replacements. If you have hardware monitoring agents in place, you sometimes get advance warnings of impending hardware failures, allowing you to preemptively bring down a server for repairs.

VMotion lets your servers keep running right through periods of hardware maintenance by migrating virtual machines to other VMware ESX server hosts with zero downtime. Your virtual machine resource allocations are preserved when moved with VMotion, and VirtualCenter’s resource monitoring tools make it easy to identify hosts with adequate resources to receive migrated VMs and guarantee committed service levels. See the section, How VMotion Works and Best Practices for further details.

In reality, VMotion can only happen on two identical physical servers, which means the VMware ESX server that applies the VMotion feature is for redundancy purposes only.

The three scenarios illustrated in the Implementation Guidelines section of this paper introduce the concept of VMotion traffic flowing in an enterprise data center network.

**VMotion Traffic Across the Access and Core Layers**

As illustrated in Figure 10, we show redundant VMware ESX servers connecting to different TOR virtual chassis, #1 and #2. The VMotion traffic, represented by the orange arrow, flows across the EX4200 switches in the same VLAN through the access layer (EX4200 access switches) and the core layer (EX8200 switches or MX960 routers).

![Figure 10: VMotion traffic flows across the access layer (Virtual Chassis) and core layer](image)

**VMotion Traffic in the Access Layer Within the Same Rack (128 Gbps Virtual Backplane)**

In the following scenarios, we illustrate the design options and advantages using the EX4200 access switch (Virtual Chassis). In this case, the redundant VMware ESX server pairs interconnect to the same Virtual Chassis configuration through the 128 Gbps virtual backplane, as shown in Figure 11. In this design option, the VMotion traffic, represented by the orange arrow, flows only a short distance across the two different racks through the 128 Gbps Virtual Chassis backplane cable.

![Diagram of VMotion traffic in the access layer within the same rack](image)
VMotion Traffic in the Access Layer Across Multiple Racks (10GbE Uplink)

In the following scenario (Figure 12), the redundant VMware server pairs in Rack 1 and Rack 2 connect to different Virtual Chassis systems, Rack 1 and Rack 2, connect to different Virtual Chassis racks through a 10GbE uplink interface. In this design, the VMotion traffic, represented by the orange arrow, flows a distance of more than 3 meters across two different racks, within the same data center through the 10GbE uplink interface with a fiber connection.

How VMotion Works and Best Practices

Let’s explore how VMotion works in more detail. First, there are several configuration requirements that must be followed for VMotion to function properly:

- VMotion is only supported by the ESX server hosts under VirtualCenter management and has to be communicating in the same Layer 2 domain.
- A dedicated Gigabit Ethernet network segment is required between the ESX server hosts to accommodate rapid data transfers.
- The ESX server hosts must share storage logical unit numbers (LUNs) on the same Storage Area Network (SAN) along with the virtual disk files for the virtual machines. For migration to occur, the ESX server host must also be contained in those shared LUNs (see Figure 13).
- The processors on the ESX server hosts must be of the same type. For example, VMotion from a Xeon host to an Opteron host is not supported because the processor architectures are quite different.
1. We start with virtual machine A running on host ESX01. We want to move VM A to our second host, ESX02, so that we can perform maintenance on host ESX0. However, VM A has active user connections and network sessions that we wish to preserve.

2. The VMotion migration is initiated from the VirtualCenter client or a VirtualCenter scheduled task or VMware SDK script. The first step is to copy the VM A configuration file to host ESX02 to establish an instance of the VM on the new host. The virtual machine configuration file is simply a small text file listing the virtual machine’s properties.

3. Next, the memory image of VM A is copied to the target host running ESX02. The memory image can be quite large, so the dedicated Gigabit Ethernet network required by VMotion allows that copy to proceed at high speed. Immediately before the VM A memory image copy begins, VMotion redirects new memory write operations on host ESX01 to a separate location on ESX01. It also keeps track of these locations through a memory bitmap which will record all VM A memory updates that occur until the virtual machine is suspended on ESX01. In that way, the full memory image is read-only and static during the VMotion operation.

4. Because the virtual disk file for VM A is stored on a Virtual Machine File System (VMFS)-formatted SAN LUN mounted by both ESX01 and ESX02, we do not need to transfer that potentially large file. VMFS is VMware’s default storage system for virtual machine files on physical Small Computer System Interface (SCSI) disks and partitions. The multiple access feature of the VMFS file system enables this time-saving method.

5. VMotion now suspends VM A on the ESX01 and copies the memory bitmap to ESX02. This step is the only one in which activity is interrupted, and that interruption is too short to cause connections to be dropped or users to notice. As soon as the memory bitmap is copied over, VMotion resumes VM A on its new home, ESX02.

6. VMotion sends a Reverse Address Resolution Protocol (RARP) packet to the production network EX4200 switch to inform it that the switch port to use for VM A has changed to the interface connecting to host ESX02. The EX4200 switch then updates its MAC table with the new MAC address and physical interface association. That preserves all network connections in the same VMotion Layer 2 broadcast domain to VM A. This step can be validated on the EX4200 switch through command-line interface (CLI) command: `show ethernet-switching table`.

7. Some modified memory pages might still reside on ESX01 after VM A has resumed—their locations being specified by the copied over memory bitmap. When VM A needs access to those pages, VMotion will “demand page” them over to ESX02.

8. VMotion completes the memory image transfer by background paging the remaining memory of VM A over to target host ESX02 and does a final commit of all modified memory pages to the full VM A memory image. Now VM A is back to using its full memory image in read/write mode.

9. The VMotion migration is now complete and we finish with a cleanup operation by deleting VM A from host ESX02.
Implementation

In this section, we illustrate and describe best practices for interconnecting VMware’s ESX 3.5 server to the data center infrastructure using two EX4200 switches with Virtual Chassis technology on each rack as “top-of-rack” switches. In this implementation, each Virtual Chassis connects to the data center’s core/aggregation layer (tier) that consists of EX8200 switches or MX960 routers (see Figure 14). Depending on your requirement for VMotion traffic flow, the implementation steps for the following three scenarios, as previously discussed in the Design Considerations section, are as follows:

- Configuring VMware connectivity to the access layer interconnecting to the core layer
- Configuring VMware connectivity at the access layer within the same rack (128 Gbps virtual backplane)
- Configuring VMware connectivity at the access layer across multiple racks (10GbE uplink)

The implementation guidelines for each of the scenarios consist of the following three major configuration steps:

1. Configuring the VMware ESX server
2. Configuring the EX4200 access switch
3. Configuring the EX8200 Ethernet Switches or MX960 Ethernet Services Router

Figure 14: Interconnecting VMware ESX server, EX4200 switch, EX8200 switch or MX960 Ethernet Services Router

Configuring VMware Connectivity to the Access Layer Interconnecting to the Core Layer

When connecting VMware’s ESX servers to EX4200 and EX8200 switches or MX960 routers, you need the following hardware components:

- Four EX4200 switches with Virtual Chassis technology/configuration
- Two EX8200 switches or MX960 routers
- Two VMware ESX servers.
- VMware ESX servers with VMotion licenses
Data Center Access Layer and Core Layer Connection Recommendations

As an example, Figure 15 illustrates the connection principles for connecting an EX4200 Virtual Chassis configuration to an EX8200 switch or MX960 router. The EX8200/MX960 connects to two Virtual Chassis clusters with the EX8200/MX960 acting as the hub between the two Virtual Chassis configurations. If communication is required between the two Virtual Chassis configurations, then the traffic will traverse the EX8200/MX960.

The following bulleted list represents the VLAN and IP addresses that Juniper Networks deployed in the solutions engineering lab, represented in Figure 15.

- **Virtual Machine Traffic**: VLAN 71, VLAN 72 (Virtual Switch Tagging)
- **VMotion Traffic**: VLAN 73 (EX Series Switch Tagging)
- **Iperf Client**: VLAN 75
- **Payload**: VLAN 71, VLAN 72
- **Client**: VLAN 75
- **EX Series switch mgmt**: 192.168.6.0/24
- **VMware server mgmt**: 192.168.3.0/24

Figure 15: Connecting VMware ESX server to access and core layer

Copyright © 2009, Juniper Networks, Inc.
VMware Server Connection Recommendations

- Bind two or more port groups to one or more NICs installed on each VMware ESX server. Connect across to different EX4200 Virtual Chassis configurations, as illustrated in Figure 16.
- Connect each virtual switch port group to one or more virtual machine guests.

The following list the major configuration procedures:
1. Configuring the EX4200 access switches
2. Configuring the EX8200 switch or MX960 core router
3. Configuring VMware server network connectivity

Configuring EX4200 Access Switches

To configure the EX4200 switch, we assume that VLAN tagging happens in virtual switches for the NIC teaming scenario. See Table 1 which explains in detail NIC teaming and VLAN tagging.

To configure EX4200 switch trunk ports (for example, ge-0/0/12 to VMware server connections and xe-0/1/0 to core layer) on one of the EX4200 Virtual Chassis configurations for NIC teaming, set the VLANs and interfaces as follows:

```
[edit]
1. set vlans vlan71 vlan-id 71
2. set vlans vlan72 vlan-id 72
3. set interfaces ge-0/0/12 description nic-teaming-port
4. set interfaces ge-0/0/12 unit 0 family ethernet-switching port-mode trunk
5. set interfaces ge-0/0/12 unit 0 family ethernet-switching vlan members [vlan71 vlan72]
6. set interfaces xe-0/1/0 description trunk-to-core
```
7. set interfaces xe-0/1/0 unit 0 family ethernet-switching port-mode trunk
8. set interfaces xe-0/1/0 unit 0 family ethernet-switching vlan members [vlan71 vlan72]

NOTE: VLAN membership can also be configured under the VLAN stanza. Either method is acceptable and there is no difference between the two. From a configuration management perspective, Juniper recommends configuring VLAN membership for access ports under the VLAN stanza and trunks ports under the interfaces (as shown above).

For more information about configuring EX Series interfaces, see Complete Software Guide for Junos for EX4200 Software.

Configuring the EX8200 Ethernet Switches

The modular EX8200 Ethernet switches are high-density, line-rate, power-efficient platforms designed for demanding data center, cloud computing and Internet exchange environments. The switches, including the eight-slot EX8208 and the 16-slot EX8216, offer high GbE and 10GbE port densities, where the ports can be segregated into separate L2 domains or VLANs. The EX8200 switches support up to 4096 VLANs. In addition to separating L2 domains, the EX8200 switches are capable of routing traffic in and out of the VLANs through their logical L3 interface, RVI (routed virtual interface).

VLAN configuration and membership follows the same structure as the EX4200 switch. The below is an example of configuring a RVI.

```
{master}[edit]
root@EX8200# set interfaces vlan unit 71 family inet address 172.16.56.2/24
root@EX8200# set vlans VLAN72 l3-interface vlan.71
```

For more information about configuring EX Series interfaces, see Complete Software Guide for Junos for EX8200 Software.

NOTE: For an EX8200 configuration example, see Appendix A, EX8200 Configuration.

Configuring the MX960 Ethernet Services Router

On the MX960, you can configure one or more bridge domains to perform Layer 2 bridging. A bridge domain is a set of logical ports that share the same flooding or broadcast characteristics. Like a VLAN, a bridge domain spans one or more ports across multiple devices. Thus, MX Series routers can function as Layer 2 switches, each with multiple bridging or broadcast domains that participate in the same Layer 2 network. You can also configure Layer 3 routing support for a bridge domain. Integrated Routing and Bridging (IRB) provides simultaneous support for Layer 2 bridging and Layer 3 IP routing on the same interface. IRB enables you to route packets to another routing interface or to another bridge domain that has a Layer 3 protocol configured. You configure a logical routing interface by including the irb statement at [edit interfaces] hierarchy level and include that interface in the bridge domain.

For more information about how to configure a routing interface, see the Junos Network Interfaces Configuration Guide.

NOTE: For an example of the software code configuration for the MX960, see Appendix A, MX960 Configuration.

Configuring the Virtual Switch on VMware ESX

Refer to Figure 17 which shows the way EX4200 switches with Virtual Chassis technology are configured on the VMware ESX virtual server and review the general recommended steps that follow.
Figure 17: VMware ESX virtual switch configuration

General Recommended Steps for Configuring a Virtual Switch on VMware ESX

1. Bind two or more port groups to one or more NICs installed on each VMware server. Connect across to different EX4200 switches with Virtual Chassis technology as illustrated in the Figure 16 and Figure 17.

2. Connect each virtual switch port group to one or more VM guest, as illustrated in Figure 17.

Configuring VMware Connectivity at the Access Layer Within the Same Rack (128 Gbps Virtual Chassis Backplane)

In this scenario, a network administrator can connect individual EX4200 switches together to form one unit and manage the unit as a single chassis, called a Virtual Chassis configuration. Up to ten EX4200 switches can be interconnected, providing a maximum of 480 Gigabit Ethernet ports. The port density increases as you expand the Virtual Chassis configuration.

To connect the VMware ESX server to the access layer over the EX4200 128 Gbps Virtual Chassis backplane, the following VLAN and NIC considerations were deployed in the Juniper Networks solutions engineering lab (see Figure 18).

- VMs in VLANs 71 and 72 benefit from the VM NIC3 and VM NIC4 bond. The EX4200 Virtual Chassis configuration load balances egress traffic across the bonded virtual machine NICs through the hash of the source virtual NIC MAC address or a hash of the source and destination IP addresses. (For IP load balancing, remember to use IP hash, configure a link aggregation group (LAG), and turn off Link Aggregation Control Protocol (LACP).)

- The EX4200 switch with Virtual Chassis technology uses all virtual machine NICs in the bond. If a link failure occurs, the EX4200 Virtual Chassis reassigns VM traffic to the remaining functional interfaces defined in the bond.

To take advantage of the higher Virtual Chassis bandwidth capacity and software redundancy features, interconnect at least two EX4200 switches in a Virtual Chassis configuration. Begin with a default configuration, consisting of two EX4200 member switches interconnected with the two dedicated 64 Gbps Virtual Chassis ports, located at the rear panel, as shown in Figure 18.
General Recommended Steps for Configuring the EX4200 Access Port for VMotion

Now that you have made the physical connectivity between VMware ESX and EX4200 access switches, you are ready to configure the VMotion interfaces (ge-0/0/1 in this case) and add the VMotion VLAN into the trunk port to the core layer (xe-0/1/0) by performing the following steps on both EX4200 switches with Virtual Chassis technology.

```
[edit]
1. set vlans vlan73 vlan-id 73 interface ge-0/0/1.0
2. set interfaces xe-0/1/0 unit 0 family ethernet-switching vlan members vlan73
3. set interfaces ge-0/0/1 description vmotion-port
4. set interfaces ge-0/0/1 unit 0 family ethernet-switching port-mode access (Note: This command is optional. All ports on the EX Series switches are access-ports by default)
```

Begin with a default configuration, consisting of two EX4200 member switches interconnected with the two dedicated 64-Gbps Virtual Chassis ports, located on the rear panel as illustrated in Figure 17. This figure depicts how to connect the two switches together using the Virtual Chassis ports.

As the Virtual Chassis configuration expands, Virtual Chassis port connectivity changes. The topology of the Virtual Chassis ports varies based upon the requirements of the network administrator. For a detailed explanation, refer to part 6 in the document Complete Software Guide for Junos for EX4200 Software.

Configuring VMware Connectivity at the Access Layer Across Multiple Racks (10GbE Uplink)

The EX4200 Virtual Chassis technology can also be extended across multiple racks, buildings, or sites through the GbE or 10GbE uplink modules. In our example we are using a 10GbE uplink (EX-UM-2XFP), as shown in Figure 19.
General Recommended Steps for Configuring Uplink for Virtual Chassis Technology

To use the uplink ports for interconnecting member switches, you must first convert the uplink ports from networking/host ports to Virtual Chassis Port Extension (VCPe). This includes configuring the uplink ports of a standalone EX4200 switch as VCPe prior to interconnecting the new member switch with the existing Virtual Chassis configuration. See Figure 20.

Figure 19: VMware connectivity across multiple racks (using 10GbE uplink)

Figure 20: Virtual Chassis racks across 10GbE
NOTE: For redundancy purposes, as shown in Figure 20, configure an uplink Virtual Chassis port in both switches SWA-0 and SWA-1. This example omits the specification of the member member-id option in configuring the uplink for SWA-0. The command applies by default to the switch where it is executed.

To configure a Virtual Chassis arrangement across multiple switches, perform the following steps:

1. Power-up and configure the mastership priority of SWA-0 (member 0) and back-up RE to be the highest possible value (255). The default mastership-priority is 128. This ensures that it functions as the primary of the expanded Virtual Chassis configuration. Refer to Figure 20 when reviewing this example.

   ```
   user@SWA-0# set virtual-chassis member 0 mastership-priority 255
   user@SWA-0# set virtual-chassis member 2 mastership-priority 255
   ```

   **Note:** The backup RE mastership priority should match the master RE to avoid preemption. For more information on Virtual Chassis best practice, please reference the Virtual Chassis Technology Best Practices implementation guide.

2. Connect the second member, SWA-1, to member 0 via the dedicated Virtual Chassis Ports (VCP).

3. Power-up the second member and verify second member joins the virtual-chassis.

   ```
   user@SWA-0> show virtual-chassis status
   ```

4. Convert the uplinks on member 0 and member 1, xe-0/1/0 and xe-1/1/0 respectively, to VCPe.

   ```
   user@SWA-0> request virtual-chassis vc-port set pic-slot 1 port 0
   user@SWA-0> request virtual-chassis vc-port set pic-slot 1 port 0 member 1
   ```

5. In preparation of extending the Virtual Chassis from Rack A to Rack B, first, without connecting the switches, boot-up both devices in Rack B, member 2 and member 3 as standalone switches. If they are not in factory-default mode, then reset them to factory default via the LCD. For more information on LCD operation, then please refer to the Hardware Guide for EX4200 Switches.

6. Next convert the uplinks on both switches in Rack B to VCPe. The command needs to be applied to both switches.

   ```
   root@> request virtual-chassis vc-port set pic-slot 1 port 0
   ```

7. Gracefully power-down both switches in Rack B.

   ```
   root@> request system halt
   ```

8. Complete the Virtual Chassis connection as shown in Figure 20.

9. Power-up member 2 switch in Rack B and verify member 2 joins the Virtual Chassis configuration.

10. Power-up member 3 switch in Rack B and verify member 3 joins the Virtual Chassis configuration.

**Summary**

Juniper Networks offers a unique value proposition for deploying server virtualization using VMware. The EX Series Virtual Chassis technology-based access layer can aggregate servers across multiple rows and multiple racks without deploying the Spanning Tree Protocol. This permits the use of superior Layer 3 connectivity (at no extra cost) from the access to the core/aggregation layer. We don’t know of any other vendor’s solution that can provide this same benefit.

By following Juniper’s recommended implementation guidelines addressed in this paper, successful server virtualization can be accomplished in a cost-effective manner while still maintaining a highly scalable and reliable network architecture. Our suggested designs establish a complete solution for server virtualization by integrating VMware’s ESX 3.5 infrastructure into the enterprise network in conjunction with deploying EX4200 switches with Virtual Chassis technology and EX8200 Ethernet switches or MX960 Ethernet Services Routers (all running on Junos Software) based on the Juniper Networks reference architecture.
Juniper Networks data center reference architecture simplifies the data center network infrastructure and provides a highly available and high-performance design that supports current as well as future applications while reducing risk and total cost of ownership.

Appendix A: Code Samples
The following code samples include configurations for enabling spanning tree on EX4200 switches with Virtual Chassis technology and for placing interfaces into a bridge domain on the MX960.

EX4200 Spanning Tree Configuration
The following configuration displays the configuration elements for enabling spanning tree on the EX4200 switch.

```
[edit protocols rstp]
interface ge-0/0/0.0 {
  cost 29999;
}
interface xe-0/1/0.0 {
  cost 29999;
}
```

EX8200 Configuration
The following configuration represents how to create Layer 2 switching interfaces on an EX8200 switch.

```
chassis {
  aggregated-devices {
    ethernet {
      device-count 1;
    }
  }
  interfaces {
    xe-0/0/0 {
      ether-options {
        802.3ad ae0;
      }
    }
    xe-0/0/6 {
      unit 0 {
        family ethernet-switching {
          port-mode trunk;
          vlan {
            members 71-72;
          }
        }
      }
    }
    xe-1/0/0 {
      ether-options {
        802.3ad ae0;
      }
    }
    ae0 {
      unit 0 {
        family ethernet-switching {
          port-mode trunk;
          vlan {
            members 71-72;
          }
        }
      }
    }
  }
```

MX960 Configuration

The following configuration represents how to create Layer 2 switching interfaces on an MX960 router.

```
[edit interfaces]
+  xe-10/1/0 {
+    flexible-vlan-tagging;
+    encapsulation flexible-ethernet-services;
+    unit 71 {
+      encapsulation vlan-bridge;
+      vlan-id 71;
+    }
+    unit 72 {
+      encapsulation vlan-bridge;
+      vlan-id 72;
+    }
```
[edit interfaces ae0]
+ unit 71 {
+   encapsulation vlan-bridge;
+   vlan-id 71;
+ }
+ unit 72 {
+   encapsulation vlan-bridge;
+   vlan-id 72;
+ }

[edit interfaces irb]
+ unit 71 {
+   family inet {
+     address 172.16.56.2/24 {
+       vrrp-group 1 {
+         virtual-address 172.16.56.1;
+         priority 190;
+         preempt;
+         accept-data;
+       }
+     }
+   }
+   family inet {
+     address 172.16.57.2/24 {
+       vrrp-group 1 {
+         virtual-address 172.16.57.1;
+         priority 190;
+         preempt;
+         accept-data;
+       }
+     }
+   }
+ }
+ unit 72 {
+   family inet {
+     address 172.16.56.2/24 {
+       vrrp-group 1 {
+         virtual-address 172.16.56.1;
+         priority 190;
+         preempt;
+         accept-data;
+       }
+     }
+   }
+   family inet {
+     address 172.16.57.2/24 {
+       vrrp-group 1 {
+         virtual-address 172.16.57.1;
+         priority 190;
+         preempt;
+         accept-data;
+       }
+     }
+   }
+ }

[edit bridge-domains]
+ VLAN71 {
+   domain-type bridge;
+   vlan-id 71;
+   interface xe-10/1/0.71;
+   interface ae0.71;
+   routing-interface irb.71;
+ }
+ VLAN72 {
+   domain-type bridge;
+   vlan-id 72;
+   interface xe-10/1/0.72;
+   interface ae0.72;
+   routing-interface irb.72;
+ }
## Appendix B: VMware Key Concepts and Terms

The following table alphabetically lists and defines the key virtual components that comprise VMware’s ESX 3.5 platform.

For further details concerning VMware concepts, products and information, see VMware Virtual Networking Concepts and visit www.vmware.com/support/pubs/vi_pubs.html.

<table>
<thead>
<tr>
<th>TERM/CONCEPT</th>
<th>DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load balancing</td>
<td>Load balancing allows you to spread network traffic from virtual machines on a virtual switch across two or more physical Ethernet adapters, giving higher throughput than a single physical adapter could provide. When you set NIC teaming policies, you have many options for creating load balancing. See VMware Virtual Networking Concepts for these load balancing options.</td>
</tr>
<tr>
<td>NIC teaming</td>
<td>You can connect a single virtual switch to multiple physical Ethernet adapters using the VMware infrastructure feature called NIC teaming. A team can share the load of traffic between physical and virtual networks among some or all of its members, and it can provide passive failover in the event of a hardware failure or a network outage. You can set NIC teaming policies at the port group level. For further details concerning how to perform an EX Series trunk configuration for NIC teaming, see Configuring the EX4200 Access Switches.</td>
</tr>
<tr>
<td>Port group</td>
<td>A port group specifies port configuration options such as bandwidth limitations and VLAN tagging policies for each member port. Network services connect to virtual switches through port groups. Port groups define how a connection is made through the virtual switch to the network. In typical use, one or more port groups are associated with a single virtual switch. Port groups aggregate multiple ports under a common configuration and provide a stable anchor point for virtual machines connecting to labeled networks. Each port group is identified by a network label, which is unique to the current host. A VLAN ID, which restricts port group traffic to a logical Ethernet segment within the physical network, is optional.</td>
</tr>
<tr>
<td>VirtualCenter</td>
<td>VMware infrastructure 3 provides a rich set of networking capabilities that integrate well with sophisticated enterprise networks. These networking capabilities are provided by VMware ESX server and managed by VMware VirtualCenter. VMware VirtualCenter provides tools for building and maintaining your virtual network infrastructure. You can use VirtualCenter to add, delete and modify virtual switches and to configure port groups with VLANs and teaming. You can use the VirtualCenter roles feature to assign the permissions a network administrator needs to manage the virtual network. For further details, see <a href="http://www.vmware.com/vmtn/resources/826">www.vmware.com/vmtn/resources/826</a>. Also see Managing VMware VirtualCenter Roles and Permissions.</td>
</tr>
</tbody>
</table>
| Virtual Ethernet adapter | Virtual Ethernet adapters are used by the virtual machine. Although there are many flavors of virtual Ethernet adapters, they primarily have the same basic attributes:  
  - Contain their own MAC addresses and unicast/multicast broadcast filters  
  - Are strictly Layer 2 devices  

Virtualization | Virtualization is a technique for hiding the physical characteristics of computing resources from the way in which other systems, applications or end users interact with those resources. This means making a single physical resource such as a server, an operating system, an application or a storage device appear to function as multiple logical resources; or making multiple physical resources such as storage devices or servers appear as a single logical resource. Virtualization also means making one physical resource appear, with somewhat different characteristics, as one logical resource. |

Virtual machine | A virtual machine (VM) is a software implementation of a machine (computer) that executes programs like a real machine. A virtual machine can be configured with one or more virtual Ethernet adapters, each of which each has its own IP address and MAC address. As a result, virtual machines have the same properties as physical machines from a networking standpoint. In addition, virtual networks enable functionality not possible with physical networks today. |

Virtual NIC | The virtual network interface card (NIC) is a virtualized network interface that resides within the virtual machine and presents the same MAC interface that an actual physical NIC interface would provide. Multiple virtual NICs can be configured to bond with the same physical interface, allowing multiple VM guests to share that interface. Internally, a virtual NIC operates the same as a physical NIC. A virtual NIC when compared to a physical NIC has the following advantage:  
  - Provides redundancy for enhanced availability |
<table>
<thead>
<tr>
<th>TERM/CONCEPT</th>
<th>DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual ports</td>
<td>The ports on a virtual switch provide logical connection points among virtual devices and between virtual and physical devices. They can be thought of as virtual RJ-45 connectors. Each virtual switch can have up to 1,016 virtual ports, with a limit of 4,096 ports on all virtual switches on a host.</td>
</tr>
<tr>
<td>VLAN tagging</td>
<td>VLANs provide for logical groupings of stations or switch ports, allowing communications as if all stations or ports were on the same physical LAN segment. Confining broadcast traffic to a subset of the switch ports or end users saves significant amounts of network bandwidth and processor time.</td>
</tr>
<tr>
<td></td>
<td>• Virtual machine guest tagging (VGT mode)—You can install an 802.1Q VLAN trunking driver inside the virtual machine, and tags will be preserved between the virtual machine networking stack and external switch when frames are passed from or to virtual switches. The format for the header of a packet tagged in this way is shown in Figure 3. Use of this mode requires that the physical switch provide a trunk.</td>
</tr>
<tr>
<td></td>
<td>• External switch tagging (EST mode)—You can use external switches for VLAN tagging. This is similar to a physical network, and VLAN configuration is normally transparent to each individual physical server. There is no need to provide a trunk in these environments.</td>
</tr>
<tr>
<td></td>
<td>• Virtual switch tagging (VST mode)—This is the most common configuration. In this mode, you provision one port group on a virtual switch for each VLAN and then attach the virtual machine’s virtual adapter to the port group instead of connecting the adapter to the virtual switch directly. The virtual switch port group tags all outbound frames and removes tags for all inbound frames. It also ensures that frames on one VLAN do not leak into a different VLAN. Using this mode requires that the physical switch provide a trunk. See Figure 3.</td>
</tr>
<tr>
<td></td>
<td>For details on using VLANs with VMware infrastructure, see the white paper titled VMware ESX Server 3802.1Q VLAN Solutions, available from the VMTN website (<a href="http://www.vmware.com/vmtn/">www.vmware.com/vmtn/</a>). Also refer to the VMware Virtual Networking Concepts for additional information on VLANs in VMware infrastructure.</td>
</tr>
<tr>
<td>VMware ESX 3.5 server</td>
<td>The VMware ESX 3.5 server is a host operating system dedicated to supporting virtual servers or virtual machines.</td>
</tr>
<tr>
<td></td>
<td>The ESX host system kernel (VMkernel) controls access to the physical resources of the server shared by the virtual machines.</td>
</tr>
<tr>
<td></td>
<td>The ESX host system ensures that the following four primary hardware resources are available to guest virtual machines:</td>
</tr>
<tr>
<td></td>
<td>• Memory</td>
</tr>
<tr>
<td></td>
<td>• Processors</td>
</tr>
<tr>
<td></td>
<td>• Storage (local or remote)</td>
</tr>
<tr>
<td></td>
<td>• Network adapters</td>
</tr>
<tr>
<td>VMware Service Console</td>
<td>VMware Service Console provides management component interfaces to the ESX server.</td>
</tr>
<tr>
<td></td>
<td>The console is based on the Red Hat Enterprise Linux server, with unique privileges for and responsibilities to the ESX system. It also provides access to the ESX host through SSH, Telnet, HTTP and FTP, and provides authentication and system monitoring services.</td>
</tr>
<tr>
<td>VMware virtual ports</td>
<td>The ports on a virtual switch provide logical connection points among virtual devices and between virtual and physical devices. They can be thought of as virtual RJ-45 connectors. Each virtual switch can have up to 1,016 virtual ports, with a limit of 4,096 ports on all virtual switches on a host.</td>
</tr>
<tr>
<td></td>
<td>The virtual ports in an ESX server provide a rich control channel for communication with the virtual Ethernet adapters attached to them. ESX server virtual ports operate in the following manner:</td>
</tr>
<tr>
<td></td>
<td>• Know authoritatively what the configured receive filters are for virtual Ethernet adapters attached to them. This means no MAC learning is required to populate forwarding tables.</td>
</tr>
<tr>
<td></td>
<td>• Unlike physical switches, know authoritatively the “hard” configuration of the virtual Ethernet adapters attached to them. This capability makes it possible to set such policies as “guest can’t change MAC address,” because the virtual switch port can essentially know for sure what is “burned into ROM” (actually, stored in the configuration file, outside control of the guest operating system).</td>
</tr>
<tr>
<td></td>
<td>For details on learning how virtual ports work, see VMware Networking Concepts.</td>
</tr>
</tbody>
</table>
### VMware virtual switch

Virtual switches are the key networking components in VMware infrastructure 3.5. You can create up to 248 virtual switches on each ESX server 3 host. A virtual switch is “built to order” at run time from a collection of small functional units. These functional units support a modular design and make up the foundation that should be followed for future development. Additionally, this modular design is beneficial for VMware, and third party developers can easily incorporate modules to enhance the system in the future. They are:

- Core Layer 2 forwarding engine
- VLAN tagging, stripping and filtering units
- Layer 2 security, checksum and segmentation offload units

### VMotion

VMware’s VMotion technology, unique to VMware, leverages the complete virtualization of servers, storage and networking to move an entire running virtual machine instantaneously from one server to another. The entire state of a virtual machine is encapsulated by a set of files stored on shared storage, and VMware’s VMFS cluster file system allows both the source and the target VMware ESX server to access these virtual machine files concurrently. The active memory and precise execution state of a virtual machine can then be rapidly transmitted over a high speed network. Since the network is also virtualized by VMware ESX, the virtual machine retains its network identity and connections, ensuring a seamless migration process.

VMware VMotion allows network administrators to:

- Perform live migrations with zero downtime, undetectable to the user
- Continuously and automatically optimize virtual machines within resource pools
- Perform hardware maintenance without scheduling downtime and disrupting business operations
- Proactively move virtual machines away from failing or underperforming servers

For further details concerning VMotion and how it applies to the scenarios described in this paper, see the VMotion Traffic Flow Considerations section.

For details on learning how VMotion works, see [www.vmware.com/support/pubs/vi/vc/vmotion.html](http://www.vmware.com/support/pubs/vi/vc/vmotion.html).

Also see VMware’s Knocking Out Downtime with Two Punches: VMotion & VMware HA.

### About Juniper Networks

Juniper Networks, Inc. is the leader in high-performance networking. Juniper offers a high-performance network infrastructure that creates a responsive and trusted environment for accelerating the deployment of services and applications over a single network. This fuels high-performance businesses. Additional information can be found at [www.juniper.net](http://www.juniper.net).