

The Role of 10 Gigabit Ethernet in Virtualized Environments

Data Center 3.0: Solutions to Accelerate Data Center Virtualization

Introduction

Virtualization is rapidly becoming an essential tool for more fully harnessing and managing the power of today's data center servers. In only a few years, standard x86 server technology has increased in performance and density so that today, multisocket, quad-core systems with 32 or more gigabytes of memory are the norm. The combination of multicore computing and virtualization software such as VMware Virtual Infrastructure has enabled IT departments to bring server sprawl under control by running multiple independent workloads on a smaller number of servers. Today, fewer servers are required to do the same work, and their utilization levels have increased - both factors that contribute to greater energy efficiency and lower power and cooling costs.

As IT departments have discovered the benefits of server consolidation, they have also found that virtualization solves an even broader set of problems. Business continuity plans based on virtualization can make disaster-recovery solutions simple, reliable, and more cost effective. Virtual desktop environments can use centralized servers and thin clients to support large numbers of users with standard PC configurations that help to lower both capital and operating costs. Virtualization allows development, test, and production environments to coexist on the same servers, and it helps decouple application deployment from server purchasing decisions. New applications can be deployed in virtual environments and scaled on demand to accommodate the evolving needs of the business.

Demand for Bandwidth

The combination of virtualization and multicore processors has been increasing the amount of work that each server can accomplish. Hardware support for virtualization achieved through Intel Virtualization Technology (IVT) and AMD Virtualization (AMD-V) technology accelerates this trend by reducing virtualization overhead and making a greater percentage of computing cycles available for applications.

While the numbers of sockets, cores, and memory slots per server have been growing by powers of two, typical server network connectivity has increased only a gigabit at a time. Network bandwidth can become a problem as the increased capacity for a server to do work creates an ever greater demand for I/O capacity. Adding a gigabit network interface card (NIC) here and there is costly because every new NIC adds to server power consumption, cabling cost and complexity, and increases the number of access-layer switch ports that need to be purchased and managed. In addition, organizations often must purchase larger, more costly servers just to accommodate the number of expansion slots needed to support increasing network and I/O bandwidth demands.

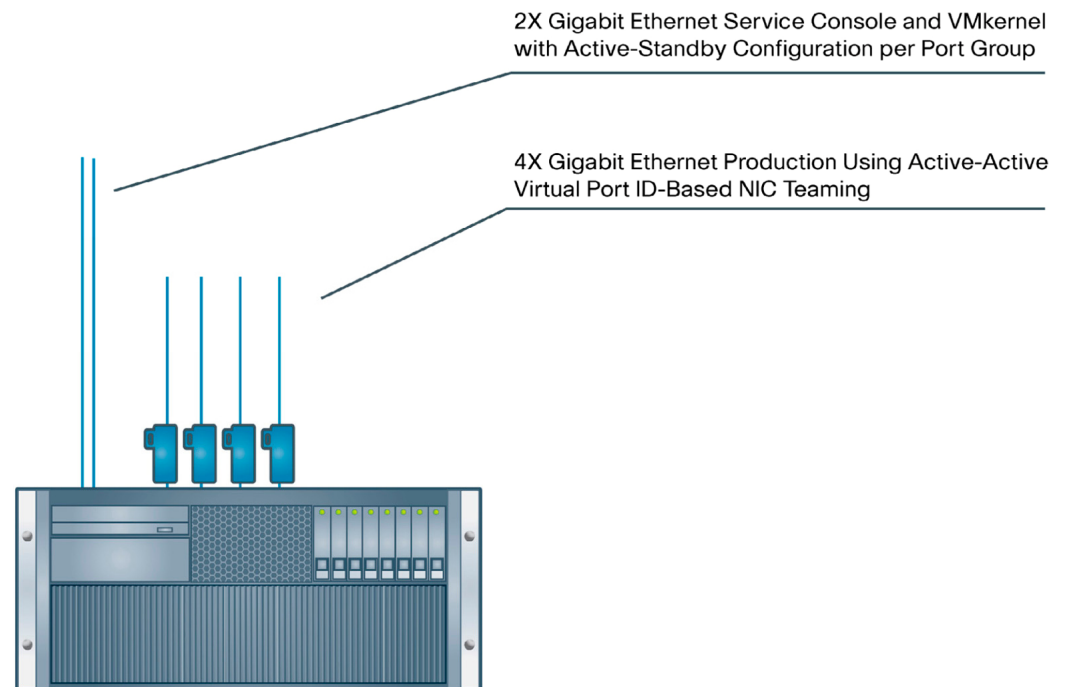
The Problem of Port Proliferation

Typical configurations for servers running VMware ESX Server use four to eight Gigabit Ethernet links, creating a need for an increased number of NICs, upstream switch ports, and cabling. It is not unusual for a server to be configured with up to eight NICs using discrete ports for each function.

Figure 1 illustrates a typical server configuration with:

- Four Gigabit Ethernet links for production traffic
- Two built-in interfaces configured in an active-standby port group for VMkernel and VMware Service Console traffic; one interface is used as the VMkernel active port and the other is used as the Service Console active port

Figure 1. Common Discrete Port Configuration for Servers Running VMware ESX Server with Separate VMkernel, Service Console, and Production Networks



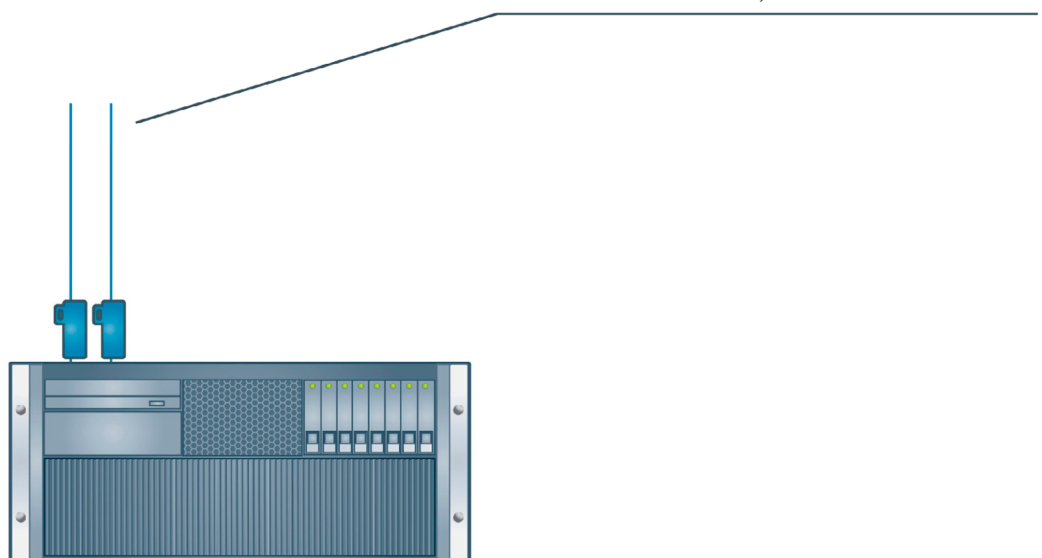
Even this 6-port configuration may be insufficient for today's workloads. Cisco's testing has shown that four virtual machines running on a two-socket, dual-core server can easily saturate four Gigabit Ethernet links, providing a hint of the bandwidth that can be required by four-socket servers equipped with quad-core processors.

The Case for 10 Gigabit Ethernet

One solution to the bandwidth problem and the proliferation of ports is to move to 10 Gigabit Ethernet technology. 10 Gigabit Ethernet delivers 10 times the bandwidth, and some switches and NICs implement additional features that support consolidation of multiple physical links onto one, including the capability to create lossless priority classes for Small Computer System Interface over IP (iSCSI), Network File System (NFS), and Fibre Channel over Ethernet (FCoE) traffic. FCoE is a straightforward encapsulation of Fibre Channel into Ethernet that preserves the lossless character and management models of native Fibre Channel networks. FCoE is supported on some 10 Gigabit

Now, rather than using four to eight Gigabit Ethernet NICs in each server, organizations can deploy just two 10 Gigabit Ethernet NICs and achieve full redundancy for availability plus additional room for expansion (Figure 2). Some Gigabit Ethernet NICs support FCoE, and all support iSCSI and NFS traffic used to access virtual disks. The benefits of 10 Gigabit Ethernet include lower cost, increased scalability, and simplified management.

2X 10 Gigabit Ethernet NICs Consolidating VMkernel, VMware Service Console, and Production Traffic



Moving to 10 Gigabit Ethernet is an economically attractive proposition that helps simplify data center infrastructure while reducing power and cooling costs.

- **Greater cost effectiveness:** As 10 Gigabit Ethernet is more widely deployed, the cost of 10 Gigabit switches and NICs has been dropping to the point where it makes economic sense to configure a pair of 10 Gigabit Ethernet NICs rather than four to eight Gigabit Ethernet NICs.
- **Fewer access-layer switches:** Dramatically reducing the number of Gigabit Ethernet ports reduces the number of NICs, upstream switch ports, and cables, simplifying data center network infrastructure while reducing capital and operating costs.
- **Lower power and cooling costs:** Every NIC and switch port contributes to a data center's power and cooling load, and reducing the number of components helps increase data center efficiency.

10 Gigabit Ethernet provides increased headroom that gives applications room to scale. The increased bandwidth of 10 Gigabit Ethernet improves scalability for both storage traffic and production traffic to virtual machines.

- ## Simplified Management

- **Ease of management:** Consolidating a number of Gigabit Ethernet links onto a single 10 Gigabit Ethernet connection dramatically decreases the number of management points, including NICs and their firmware, cables, and switch ports.
- **Simplified VMware VMotion support:** Moving a virtual machine from one server to another requires identical I/O configurations between servers. Standardizing on only one pair of 10 Gigabit Ethernet NICs per server significantly simplifies the task of helping ensure consistent I/O configurations on servers running VMware ESX Server.
- **Flexible bandwidth allocation:** Recommended VMware configurations separate VMkernel, VMware Service Console, and production traffic. Cisco® Data Center Ethernet features allow multiple classes of traffic to share the same link without the interference that can occur with VLANs alone. This feature makes physical network separation unnecessary.
- **Fewer points of failure:** Decreasing the number of devices reduces the number of failure points in the access layer, contributing to higher availability.

Cisco Nexus™ 5000 Series Switches comprise a family of line-rate, low-latency, lossless 10 Gigabit Data Center Ethernet and FCoE switches. Together with an ecosystem of NIC and cabling partners, the Cisco Nexus 5000 Series delivers the network capacity and management flexibility needed by multicore, multisocket servers running at high utilization levels because of software such as VMware Virtual Infrastructure.

Figure 3. The Cisco Nexus 5020 Switch, the First in the Cisco Nexus 5000 Series, Provides 40 Fixed 10 Gigabit Ethernet Ports and the Capability to Support Up to Two Expansion Modules with Three Combinations of Ethernet and Fibre Channel Connectivity Available.



The switch family supports Cisco Data Center Ethernet capabilities that increase the reliability, efficiency, and scalability of Ethernet networks. These capabilities enhance the ability to aggregate multiple Gigabit Ethernet links onto a single 10 Gigabit Ethernet link, making physical network separation unnecessary. Two Data Center Ethernet features pertinent to VMware environments are Priority Flow Control (PFC) and Bandwidth Management:

- **PFC** allows eight priority classes of traffic to be managed separately on a single link. It can be used, for example, to prioritize VMware VMotion traffic and to create lossless networks for storage traffic.
- **Bandwidth Management** supports minimum bandwidth guarantees per priority class, while providing the flexibility to absorb traffic spikes by allowing utilization of unused bandwidth (Figure 4). This feature allows both VMkernel and production traffic to coexist on the same link without the risk of interference, eliminating concern that production traffic will interfere with VMware VMotion traffic on the same link and cause the link to fail.

The example in Figure 4 shows how varying traffic loads would be shaped on a 10 Gigabit Ethernet network with bandwidth guarantees of 3 Gbps each for VMkernel and storage traffic, and 4 Gbps for LAN traffic. In time period t1, the offered traffic is less than or equal to the guaranteed minimums. During period t2, LAN traffic increases to 4 Gbps, filling its minimum. During time period t3, the offered VMkernel traffic drops to 2 Gbps and LAN traffic rises to 6 Gbps. The flexible bandwidth management mechanism allows LAN traffic to increase to 5 Gbps, using all available bandwidth but not interfering with VMkernel and storage traffic that is at or below their minimum guaranteed bandwidth amounts.

The diagram illustrates traffic utilization on a 10 Gigabit Ethernet link across three time periods: t1, t2, and t3. It shows how different types of traffic (VMkernel, Storage, and LAN) are prioritized and how their utilization changes as offered traffic increases.

Offered Traffic:

- Scenario 1 (Green):** 3 Gbps at t1, 3 Gbps at t2, 2 Gbps at t3.
- Scenario 2 (Yellow):** 3 Gbps at t1, 3 Gbps at t2, 3 Gbps at t3.
- Scenario 3 (Blue):** 3 Gbps at t1, 4 Gbps at t2, 6 Gbps at t3.

Traffic Utilization Achieved on 10 Gigabit Ethernet Link:

- Scenario 1 (Green):** VMkernel Traffic (3 Gbps) is utilized at t1 and t2. Storage Traffic (3 Gbps) is utilized at t2. LAN Traffic (4 Gbps) is utilized at t3.
- Scenario 2 (Yellow):** VMkernel Traffic (3 Gbps) is utilized at t1 and t2. Storage Traffic (3 Gbps) is utilized at t2. LAN Traffic (4 Gbps) is utilized at t3.
- Scenario 3 (Blue):** VMkernel Traffic (3 Gbps) is utilized at t1 and t2. Storage Traffic (3 Gbps) is utilized at t2. LAN Traffic (4 Gbps) is utilized at t3.

The diagram shows that as offered traffic increases, the utilization of different traffic types changes. For example, in Scenario 3, the offered traffic at t3 is 6 Gbps, but the utilization is limited to 5 Gbps (3 Gbps for VMkernel, 3 Gbps for Storage, and 4 Gbps for LAN, with 1 Gbps of LAN traffic being dropped).

Migrating VMware Environments to 10 Gigabit Ethernet

- Select a 10 Gigabit Ethernet NIC from one of Cisco's partners. If you wish to consolidate I/O using FCoE, use a converged network adapter that is supported by VMware ESX Server.
- Use a dual-port NIC or two NICs depending on your availability requirements. Establish high-availability connectivity to two upstream switches using active-standby and active-active vSwitch and VMNIC configurations as illustrated in Figure 5.
 - Use an active-standby configuration for the VMware Service Console, where VMNIC0 is active and VMNIC1 is standby.
 - Use an active-standby configuration for VMkernel traffic, where VMNIC1 is active and VMNIC0 is standby.
 - Use active-active virtual port ID-based NIC teaming for production traffic.
- Use VLAN tagging to securely separate VMware Service Console, VMkernel, and production traffic.

The diagram illustrates an Active-Active Network Configuration for VMware ESX Server Hosts, showing two hosts connected to two Cisco Nexus 5000 Series Switches via Production VLANs and Service Console VMkernel VLANs.

Legend:

- Orange line: Active Layer 2 Links for VMware Service Console and VMkernel Traffic
- Gray line: Standby Links for VMware Service Console and VMkernel Traffic
- Blue line: VMkernel Traffic Active Layer 2 Links for Virtual Machine Production Traffic

Configuration Details:

- VMware ESX Server Hosts:** Two hosts are shown, each with a vSwitch and two VMkernel interfaces (VMNIC0 and VMNIC1).
- VMkernel Traffic:** Active Layer 2 Links (Blue) connect VMNIC0 to the left switch and VMNIC1 to the right switch. Standby Links (Gray) connect VMNIC1 to the left switch and VMNIC0 to the right switch.
- Service Console and VMkernel Traffic:** Active Layer 2 Links (Orange) connect VMNIC0 to the right switch and VMNIC1 to the left switch. Standby Links (Gray) connect VMNIC1 to the right switch and VMNIC0 to the left switch.
- Virtual Machines:** VM1 and VM2 are connected to the left vSwitch via Port Group 1. The Service Console and VMkernel interfaces are connected to the right vSwitch via Port Group 1.
- Network Topology:** The two switches are connected to each other via a Service Console VMkernel VLAN. The two hosts are connected to each other via a Production VLAN.

Active-Active Configuration:

- Active-Active VMNIC0/VMNIC1:** This configuration is shown for the left host, where both VMNIC0 and VMNIC1 are active and connected to the left switch.
- Active-Standby VMNIC0/VMNIC1:** This configuration is shown for the right host, where VMNIC0 is active and VMNIC1 is in standby mode.

- For FCoE storage traffic, use PFC to establish a lossless class, and further increase efficiency by using jumbo frames, as supported by Cisco Nexus 5000 Series Switches. For iSCSI and NFS traffic, use jumbo frames to increase efficiency when supported by VMware, a feature targeted for VMware ESX Server 4.0.
- Use bandwidth management features to manage the way that offered loads from the VMware Service Console, VMkernel, and production traffic classes are shaped onto the 10 Gigabit Ethernet link. Use rate limiting in the vSwitch or use the NIC's I/O virtualization (IOV) driver, if supported, to manage bandwidth from virtual machines to the NIC.
- Enable congestion management capabilities of the Cisco Nexus 5000 Series that help prevent congestion in the aggregation layer or network core. This feature allows the switch to differentiate between short-term bursts of traffic and long-term congestion, absorbing short-term bursts without any effect on traffic.

Conclusion: Putting 10 Gigabit Ethernet to Work

Page 7

Bandwidth that supports I/O-intensive workloads is one reason for deploying 10 Gigabit Ethernet in the data center; increased responsiveness is another. Virtual machines can boot more quickly, and software running within them can be more responsive when I/O is no longer limited to 1 Gbps. 10 Gigabit Ethernet delivers I/O consolidation on a single network with 10 times the headroom compared to Gigabit Ethernet networks. These features help support the bursty nature of applications in enterprise environments. The compelling arguments for 10 Gigabit Ethernet in the data center make it the technology of choice for today's powerful, multicore servers running software such as VMware Virtual Infrastructure.

For more information on the Cisco Nexus 5000 Series Switches, visit <http://www.cisco.com/go/nexus5000>.

For more information on VMware products, visit: <http://www.vmware.com>.



For more information, visit: www.vmware.com



Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883



VMware, Inc
3401 Hillview Ave
Palo Alto, CA 94304
USA
www.vmware.com
Tel: 1-877-486-9273 or 650-427-5000
Fax: 650-427-5001

Copyright © 2010. VMware, Inc. All rights reserved. Protected by one or more U.S. Patent Nos. 6,397,242, 6,496,847, 6,704,925, 6,711,672, 6,725,289, 6,735,601, 6,785,886, 6,789,156, 6,795,966, 6,880,022, 6,944,699, 6,961,806, 6,961,941, 7,069,413, 7,082,598, 7,089,377, 7,111,086, 7,111,145, 7,117,481, 7,149, 843, 7,155,558, 7,222,221, 7,260,815, 7,260,820, 7,269,683, 7,275,136, 7,277,998, 7,277,999, 7,278,030, 7,281,102, 7,290,253, 7,356,679 and patents pending.

Cisco, the Cisco logo, and Cisco Systems are registered trademarks or trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries. All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company.
(0807R)

C11-496511-02 05/10