CASE STUDY
Intel® Ethernet Server Adapters
FedEx Services

FedEx Moves Data at Full Speed with Virtual Machine Device Queues

A proof of concept by Intel and FedEx Services establishes best practices for using Virtual Machine Device Queues (VMDq) technology and VMware NetQueue* to allow virtual machines to fully utilize 10 Gigabit Ethernet.

FedEx Corporation

CHALLENGE

FedEx wanted to reduce network complexity and speed up infrastructure deployment using 10 Gigabit Ethernet (10GbE) instead of 1 Gigabit Ethernet (1GbE) to support demanding data replication and distribution requirements at a new data center. FedEx deployed a VMware ESX* hypervisor solution as its virtualization solution of choice, but after initial testing were looking to improve network performance on their new 10GbE links. FedEx undertook a proof-of-concept (PoC) study with Intel to determine how to improve 10GbE network throughput to near line rate.

SOLUTION

Virtual Machine Device Queues (VMDq, a component of Intel® Virtualization Technology for Connectivity), together with VMware NetQueue*, helped FedEx better utilize 10GbE bandwidth, allowing it to meet its requirements while managing costs in its new data center. Using 10GbE to consolidate multiple Ethernet connections reduces the number of adapters and cables in the solution while improving power management and overall complexity.

INFRASTRUCTURE COMPONENTS

- Intel® Xeon® processor 5500 series
- Intel® 10 Gigabit AF DA Dual Port Server Adapter with VMDq
- VMware ESX 4.0, part of the VMware vSphere* product family
- Red Hat Enterprise Linux* 5.3 64-bit

As a global transportation, e-commerce, and business systems provider, managing large quantities of data is a core competency for FedEx. When the company started planning infrastructure for their new data center, the primary considerations included determining the most effective network connectivity, in both a technical and financial sense, to use in a VMware vSphere* virtualized environment.

Early in the design phase, FedEx network engineers compared FTP throughput in their environment using 1GbE and 10GbE server adapters. They measured rates of 320 Mbps over the 1GbE links and an increase to only 560 Mbps over the 10GbE links—an increase of only 1.75x, although the line rate was 10x higher. To address that disparity, FedEx initiated a PoC study in collaboration with Intel.

Establishing the Proof-of-Concept Approach

For the PoC, FedEx and Intel chose a file transfer workload using a variety of file-copy tools, to emulate a common class of tasks routinely performed on FedEx's production network. Engineers at the two companies hypothesized that the VMDq capability of Intel® Ethernet Server Adapters could overcome I/O bottlenecks so the test workload could achieve nearly line-rate throughput:

- **Hardware assist for traffic sorting.** VMDq enhances network data processing by offloading data sorting from the hypervisor to the network silicon. The data packets with the same destination are grouped together in a single queue and then routed to the most appropriate processor on which the destination virtual machine (VM) is running.
- **Streamlined virtualization.** With VMDq, a compatible hypervisor, such as VMware ESX, directs traffic to its destination without having to perform packet filtering, which improves overall processor utilization and throughput levels.
- **Hardware/software synergy.** VMware NetQueue is the software feature within VMware ESX that provides support for VMDq.

The lower cost of implementing 10GbE links instead of a larger number of 1GbE links was a key driver for this study. Therefore, the team also explored the use of less expensive direct attach adapters and twinaxial cabling, coupled with SFP+ modules, for short runs (the technology's maximum range is seven meters). That approach both affords lower latency and dramatically lowers power requirements. The resulting lower operating costs make twinaxial a valuable adjunct to standard 10GbE implementations.

As a relatively simple initial test case, the team began with Netperf* (a synthetic benchmark) as the workload in a non-virtualized environment. The team achieved throughput with the non-virtualized Netperf workload of almost 10 Gbps.

With the near-line-speed result using Netperf in mind, the team moved to a real application workload, still in a non-virtualized environment. Because standard file-copy tools are not well-threaded, the workload was unable to take full advantage of the parallel hardware of the processor and Ethernet adapter, and throughput was severely constrained. Unlike the Netperf case, the single data stream throughput with a real-world workload was under 2 Gbps with the majority of tools tested. Using eight simultaneous data streams (corresponding to the number of processor cores) overcame that threading limitation, achieving throughput with several tools of more than 9 Gbps.

Proof-of-Concept Results with Virtualized, Real-World Workloads

The Intel and FedEx team next turned their attention to a test case that more closely emulated the eventual production environment—real-world workloads using virtualization in a VMware vSphere environment. Building on the results so far, they recognized the value of matching the number of single-stream VMs to the number of processor cores available. Therefore, the results shown in Figure 1 reflect the use of real-world workloads with one file-transfer stream in each of eight VMs running on a dual-processor system based on the Intel Xeon processor 5500 series (quad-core).

Figure 1 shows several cases where enabling VMDq results in substantial performance increases. Moreover, enabling VMDq does not have a negative effect in any of the test cases in a virtualized environment. In the first five cases, AES 128-bit crypto processing is enabled. Crypto processing is very CPU-intensive and can limit file-transfer performance. The last three cases show that the throughput is much higher with crypto disabled.

To further illustrate the impact of VMDq on file-transfer operations, the team performed similar testing on iSCSI (Internet small computer system interface) Read operations. The results are shown in Figure 2. Here, the iSCSI target is a ramdisk device located on the source VM, while the iSCSI initiator is located on the destination VM. The VMDq benefit in performing iSCSI reads here is obvious, increasing throughput from approximately 5 Gbps on average to more than 9 Gbps in three of the six test cases.

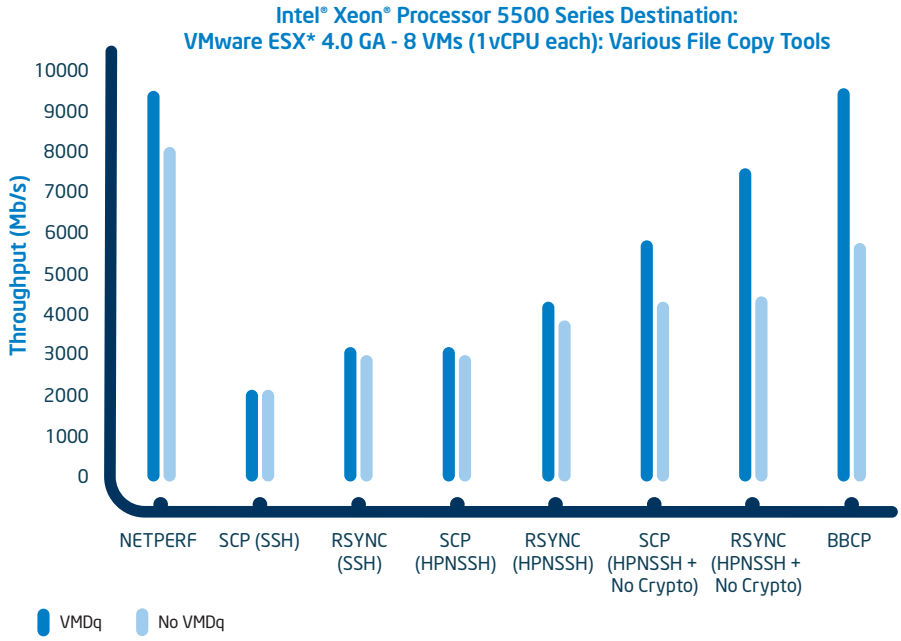

Figure 1. File-transfer throughput using various file-copy tools with and without Virtual Machine Device Queues.

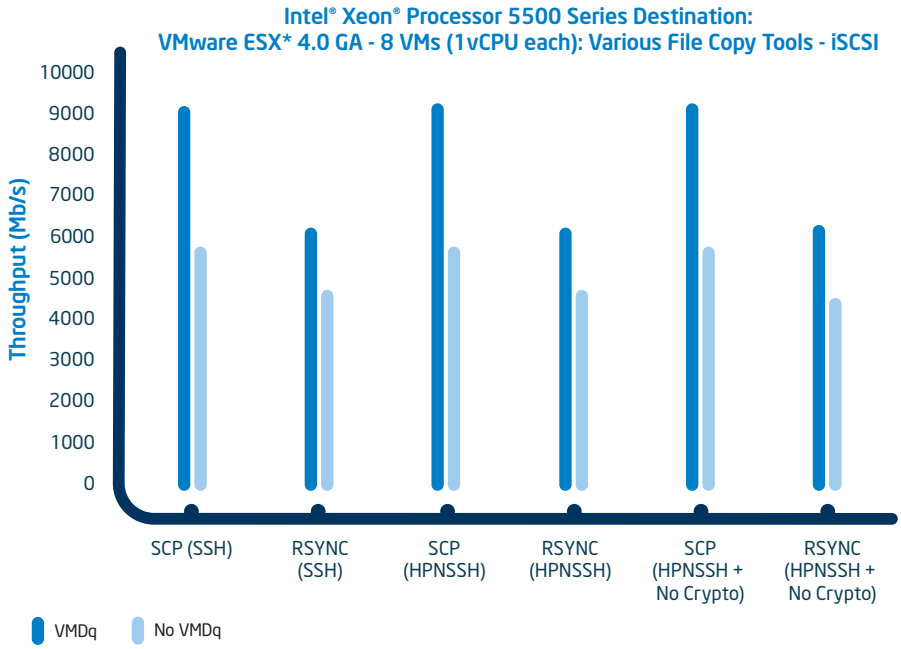

Figure 2. File-transfer throughput using various file-copy tools for iSCSI read with and without Virtual Machine Device Queues.

Best Practices and Recommendations

The core finding of this PoC study is that VMDq can, in fact, help FedEx achieve its goal of near-line-speed throughput from 10GbE. In conjunction with that finding, the Intel and FedEx PoC team offers the following best practices and recommendations:

- **Use care in choosing PCI Express* (PCIe) slots.** Check with the server manufacturer to ensure that the slot being used meets the PCIe Gen1 x8 or PCIe Gen2 x4 specifications for one 10GbE port or the PCIe Gen 2 x8 specifications for two 10GbE ports.
- **Consider the use of direct attach server adapters and twinaxial cabling.** In many cases, twinaxial can greatly reduce the cost of implementing 10GbE.
- **Verify BIOS settings.** While reporting tests for these settings is outside the scope of this paper, turning the following BIOS settings ON had positive or neutral effects: VT-x, VT-d, "energy-efficient mode" (Turbo, EIST), SMT, and NUMA.
- **Use synthetic benchmarks sparingly.** They can be useful in evaluating subsystem performance but do not necessarily approximate the behavior of real-world workloads and can be misleading in estimating throughput.
- **Choose tools and usage models carefully.** Look for limitations from tool-threading behavior, increase parallelism where possible, and be aware of potential bottlenecks from cryptography.

Get answers in The Server Room:
communities.intel.com/community/openportit/server

Learn more about Virtual Machine Device Queues (VMDq) and other Intel® I/O virtualization technologies:
www.intel.com/go/vtc

SOLUTION PROVIDED BY:

Printed in USA 0909/BY/MESH/PDF ♻ Please Recycle 322583-001US