

Best Practices for Simplifying Your Cloud Network

Maximize the networking benefit in private cloud infrastructures

Today's Intel® Ethernet 10 Gigabit Server Adapters can greatly reduce networking complexity in private cloud environments. To deliver robust and reliable cloud services, dramatic optimizations in networking can be attained by leveraging the right intelligent offloads and technologies within Intel® Ethernet devices.

Many organizations are following the lead of public cloud services such as Amazon EC2* and Microsoft Azure* to create private cloud infrastructures behind their own firewalls for internal use.

Those private cloud networks typically position IT organizations as internal service providers in their own right, delivering services within their organizations on a per-use basis. This paper provides best practices for private cloud networking, established through research and real-world implementations by Intel.

Best Practices Transform Networking to Meet Cloud Challenges

Cloud computing represents an evolution of data center virtualization, where services and data reside in shared, dynamically scalable resource pools available to any authenticated device over the Internet or within an internal network. The computing industry has begun to provide technologies for cloud services to build on core virtualization benefits including abstraction, scalability, and mobility.

Organizations that implement cloud computing can create automated, dynamically adjustable capacity on demand. The elasticity they realize as a result delivers cost and agility benefits, provided they can meet the key challenges they face during the transition. This can be accomplished in part by following a strategic approach of creating standards-based systems that are simplified, efficient, and secure.

Table of Contents

Best Practices Transform Networking to Meet Cloud Challenges	1
Simplification Meets the Challenges of Complexity and Cost.....	3
Efficiency Meets the Challenges of Scalability and Quality of Service	3
Security Meets the Challenges of Hardened Flexibility	3
Best Practice #1: Consolidate Ports and Cables Where Appropriate	4
Best Practice #2: Converge Data and Storage onto Ethernet Fabrics	5
But How will FCoE Perform Compared to Fibre Channel?	6
Best Practice #3: Maximize I/O Virtualization Performance and Flexibility	7
I/O Virtualization with Intel® VT-c	8
Best Practice #4: Enable a Solution That Works with Multiple Hypervisors	9
Best Practice #5: Utilize QoS for Multi-Tenant Networking	10
Conclusion	11

Simplification Meets the Challenges of Complexity and Cost

Building out virtualized and cloud infrastructures does not inherently reduce complexity. On the contrary, structured design practices are needed, which can eliminate the need for multiple fabrics and enable many cables and connections to be converged onto a single fabric type and potentially a single wire.

Figure 0 illustrates an initial-state environment before such practices have been implemented, where a proliferation of server network ports, cables, and switches can counter the cost and flexibility benefits organizations set out to achieve with cloud computing.

Infrastructure complexity and associated costs often arise as a result of organic development from prior topologies, and they can be readily addressed using the best practices outlined in this paper.

Efficiency Meets the Challenges of Scalability and Quality of Service

IT organizations must simultaneously meet the demands for increased capacity that their customers demand while also ensuring the level of service those customers require to ensure business continuity. Scalability and quality of service (QoS) requirements defined by network policies create new networking challenges for cloud architects to address.

These include I/O bottlenecks and contention for network resources brought about by large numbers of virtual machines (VMs) per host, as well as the need to handle usage peaks from large numbers of customers.

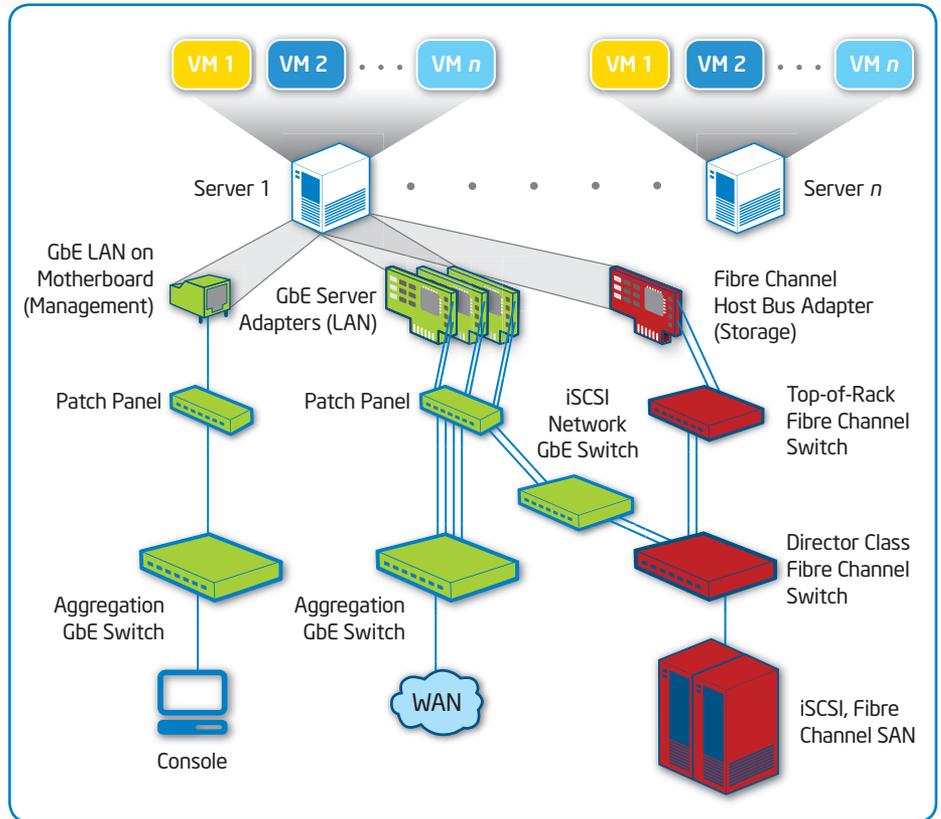


Figure 0. Many networks use Gigabit Ethernet for LAN traffic, which drives up port count, and Fibre Channel for storage traffic, which adds the complexity of another fabric. Similar issues exist with iSCSI (Internet Small Computer System Interface) SANs, as well, albeit on an Ethernet fabric.

Security Meets the Challenges of Hardened Flexibility

Moving business-critical workloads to cloud infrastructures, especially by regulated or otherwise particularly security-conscious industries, requires hardened cloud infrastructures. At the same time, they must deliver the flexibility that is at the core of agility and elasticity requirements that draw organizations to cloud computing.

Intel® Ethernet 10Gb Server Adapters support trusted, standards-based, open solutions for cloud computing:

- **Intelligent performance.** Virtualized I/O enabled by intelligent offloads
- **Flexible operation.** Broad OS support and dynamic configuration across ports and virtual machines
- **Proven reliability.** Thirty years of Ethernet products that “just work”

Best Practice #1: Consolidate Ports and Cables Where Appropriate

The increasing compute capacity of servers drives the need for larger amounts of data throughput to maximize platform-utilization levels. In organizations that have standardized on GbE server adapters, that need for connectivity and bandwidth can lead to very high physical port counts per server, and the use of 8-12 GbE server adapter ports per server is not uncommon.

In particular, many network architects continue to follow the recommendation of deploying multiple GbE physical port connections for a host of VMs on the server, as well as dedicated ports for

live migration and the service console. Provisioning high numbers of GbE server adapter ports per server increases network cost and complexity, making inefficient use of cable management and power and cooling resources in the data center.

A simple way to reduce the number of physical connections per server is to consolidate multiple GbE ports to a smaller number of 10GbE ports, in those areas where it makes sense. In many cases, network architects may determine that GbE ports are valuable in some instances and choose to maintain a combination of GbE and 10GbE ports for LAN traffic. In particular, management ports are often

kept on dedicated GbE ports, and some hypervisor vendors suggest separate physical ports for VM migration.

Figure 1 shows network connectivity analogous to that in Figure 0, but taking advantage of 10GbE to reduce the number of physical ports needed for LAN traffic. Such consolidation can dramatically reduce cabling and switch requirements for lower equipment cost, power and cooling requirements, and management complexity.

Note: Throughout this paper, the progression of numbered network diagrams builds on one another, implementing each best practice on top of the previous one.

The use of 10GbE as opposed to GbE enables the following advantages:

- **Simplicity and efficiency.** Fewer physical ports and cables make the environment simpler to manage and maintain, reducing the risk of misconfiguring while using less power.
 - **Reduced risk of device failure.** Fewer server adapters and port connections reduce the points of potential failure, for lower maintenance requirements and overall risk.
 - **Increased bandwidth resources.** The 10GbE topology increases the overall bandwidth available for each VM, better enabling bandwidth elasticity during peak workloads.
- 10GbE also takes better advantage of overall bandwidth resources because multiple logical connections per server adapter can share excess headroom in terms of throughput. Redundancy is provided by the use of two 10GbE connections on a standard two-port 10GbE server adapter.

Moreover, features that provide high performance with multi-core servers and optimizations for virtualization make 10GbE the clear connectivity medium of choice for the data center.

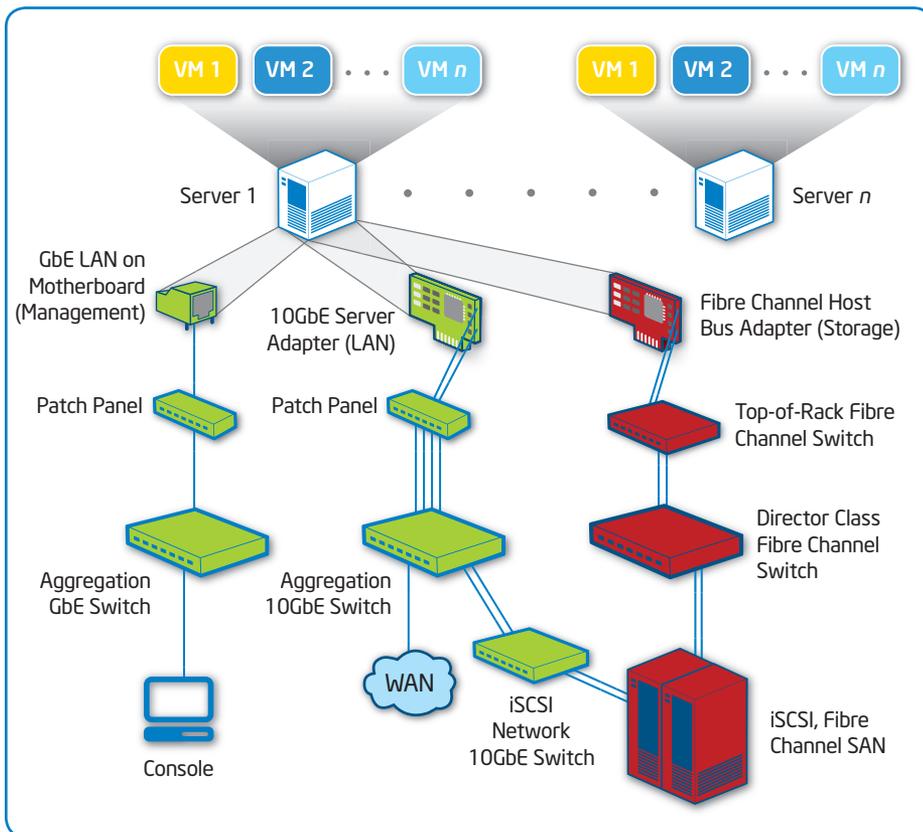


Figure 1. Consolidating Gigabit Ethernet (GbE) server adapters onto 10GbE reduces the number of ports and cables in the network, lowering complexity and costs.

Best Practice #2: Converge Data and Storage onto Ethernet Fabrics

Significant efficiencies are possible by passing storage traffic over Ethernet, as shown in Figure 2. This approach simplifies the server configurations to one fabric technology (Ethernet), providing for cost-effective, optimized transmission of all data streams—LAN, management, and storage. To support this functionality, Intel® Ethernet takes advantage of the mature trusted protocols integrated into the network devices as well as the mainstream OSs.

NAS (Network Attached Storage) and iSCSI (Internet Small Computer System Interface) SANs natively use Ethernet for transmission, while Fibre Channel SANs use a different fabric. FCoE (Fibre Channel over Ethernet) is a new technology that allows Fibre Channel traffic to be encapsulated in standard Ethernet frames that can be passed over Intel® Ethernet 10 Gigabit server adapters.

Because Fibre Channel requires guaranteed successful data delivery, FCoE requires additional Ethernet technology specifically to provide lossless transmission. Intel Ethernet provides that functionality using the IEEE 802.1 data center bridging (DCB) collection of standards-based extensions to Ethernet, which allows for the definition of different traffic classes within the 10GbE connection. In this case, storage is assigned to its own traffic class that guarantees no packet drop from the server to the storage target.

Using Ethernet fabric for storage allows the elimination of expensive, power-hungry host bus adapters (HBAs) or converged network adapters (CNAs) and associated switching infrastructure. FCoE with Intel Ethernet can help reduce both equipment costs and operating expenses associated with storage networking. Moreover, the use of only one networking fabric instead of two creates a simpler infrastructure, based on technology

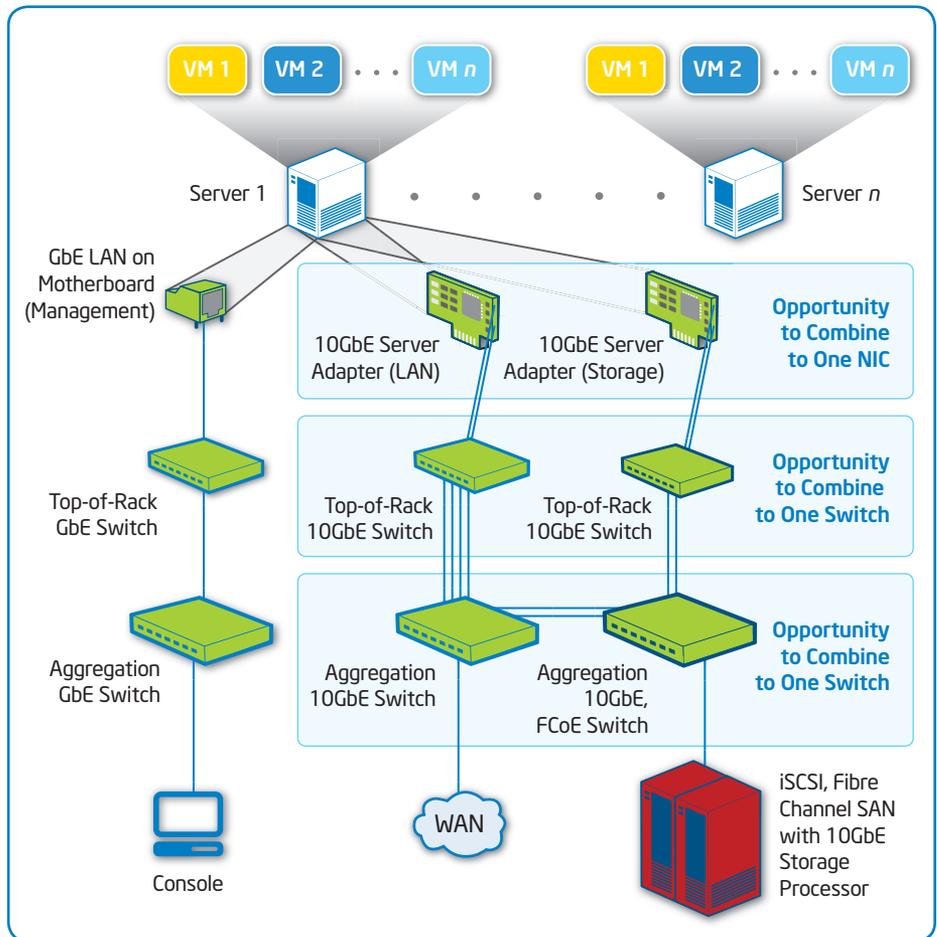


Figure 2. Intel® Ethernet 10GbE Server Adapters support storage technologies, such as NAS (Network Attached Storage), iSCSI (Internet Small Computer System Interface), and FCoE (Fibre Channel over Ethernet), enabling both storage and LAN traffic to be passed over Ethernet.

that is familiar to IT support staff. It is also compatible with standard, IP-based management technologies that are already in use within most IT organizations.

Passing storage traffic over Ethernet provides two options for network architects, depending on whether their internal procedures require dedicated links and infrastructure for storage traffic:

- **Option 1. Use dedicated 10 Gigabit Ethernet (10GbE) connections for storage (as shown in Figure 1).** Using Ethernet in place of Fibre Channel for storage traffic enables the cost and power savings already mentioned and standardization on a single type of

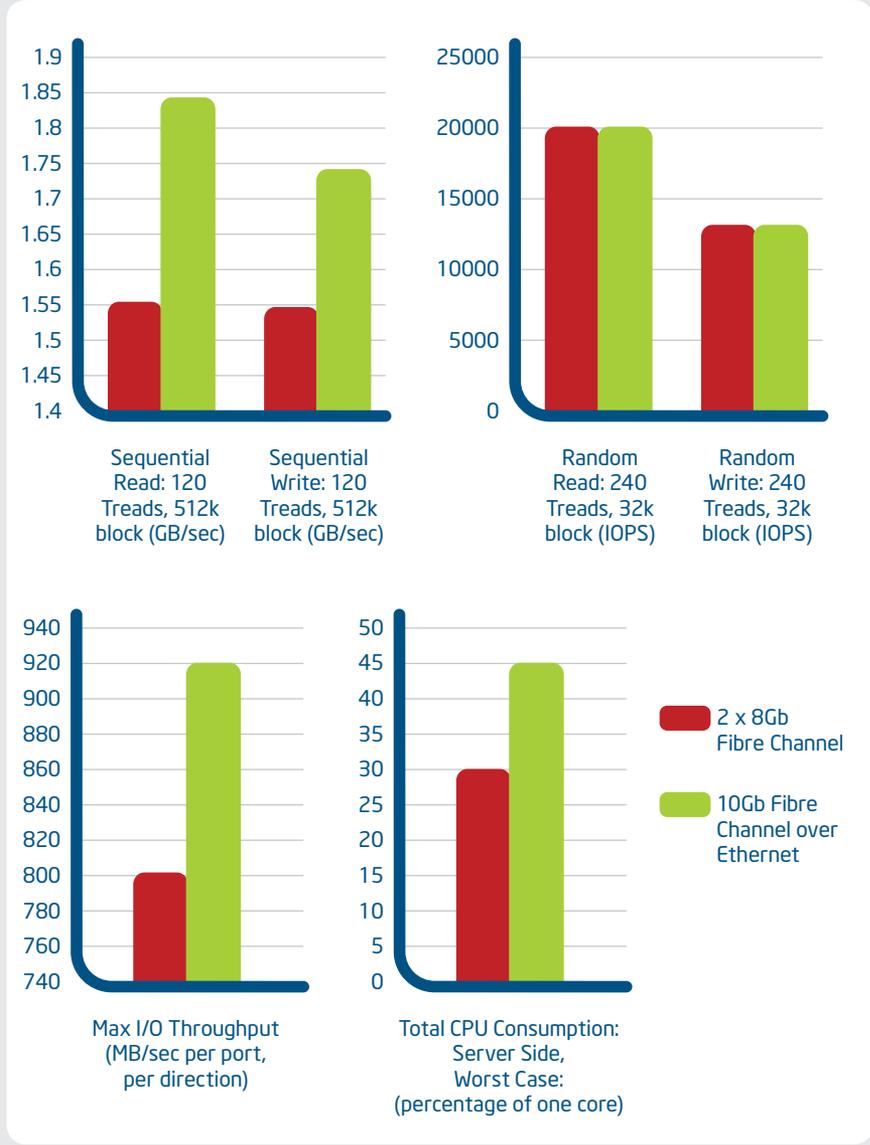
server adapter and switch. Note that a single dual-port server adapter can provide dedicated ports for both LAN and SAN traffic.

- **Option 2. Pass LAN and storage traffic over a single wire.** Building a topology where both types of traffic use the same 10GbE server adapter ports and cables provides optimal reduction in port count and associated capital and operating costs.

Many organizations adopt a phased approach, using dedicated ports to physically isolate LAN and storage traffic at first and then later reducing port count and complexity by consolidating both types of traffic onto a single interface.

BUT HOW WILL FCoE PERFORM COMPARED TO FIBRE CHANNEL?

As a proof-of-concept test, Yahoo! compared the performance of Fibre Channel over Ethernet (FCoE) networking with native Fibre Channel. The goal of this testing was to determine whether the cost and flexibility advantages of FCoE were accompanied by significant trade-offs in read/write speed or throughput. The raw performance results are summarized in Figure A:¹



The primary conclusion drawn by the team at Yahoo! is that FCoE provides competitive performance versus Fibre Channel, with comparable results in some cases and improvements in others:

- Random I/O is about the same between the two fabrics; transport capacity is not a bottleneck.
- Sequential I/O throughput is improved by 15 percent using FCoE relative to native Fibre Channel; transport capacity is a bottleneck.
- Transport capacity is improved by 15 percent using FCoE relative to native Fibre Channel.
- The increased amount of processor resources that FCoE consumes is negligible compared to that consumed by native Fibre Channel.

Figure A. Proof of Concept with Yahoo! at EMC World, 2011 showed comparable results between Fibre Channel over Ethernet versus native Fibre Channel.

Best Practice #3: Maximize I/O Virtualization Performance and Flexibility

Each generation of processors delivers dramatic compute-performance improvements over what came before, and this has driven the use of virtualization to maximize the efficient use of that compute. To maintain a balanced compute environment in a virtualized environment, it is therefore necessary to provide very high levels of network I/O throughput and performance VMs.

10GbE provides the means to increase raw bandwidth, but to use the effective bandwidth across the VMs, optimizations are required in the virtualized environment. Intel® Virtualization for Connectivity (Intel® VT-c) provides these virtualization capabilities in the Intel Ethernet silicon. Intel VT-c transitions physical network models traditionally used in data centers to more efficient models by providing port partitioning, multiple Rx/Tx queues, and on-controller QoS functionality that can be used in virtual server deployments. These functions help reduce I/O bottlenecks and improve overall server performance by offloading the data sorting and queuing functionality to the Intel Ethernet controller.

Intel VT-c provides native balanced bandwidth allocation and optimized I/O scalability. These functions help reduce I/O bottlenecks and help improve overall server performance by accelerating functionality in the Ethernet controller, as shown in Figure 3. That acceleration reduces the data processing bottleneck associated with virtualization by providing queue separation for each VM and intelligently offloading switching and sorting operations from the hypervisor to the server adapter hardware.

To improve hypervisor performance, Intel VT-c supports separate queuing for individual VMs on the network controller itself, by relieving it of the burden of

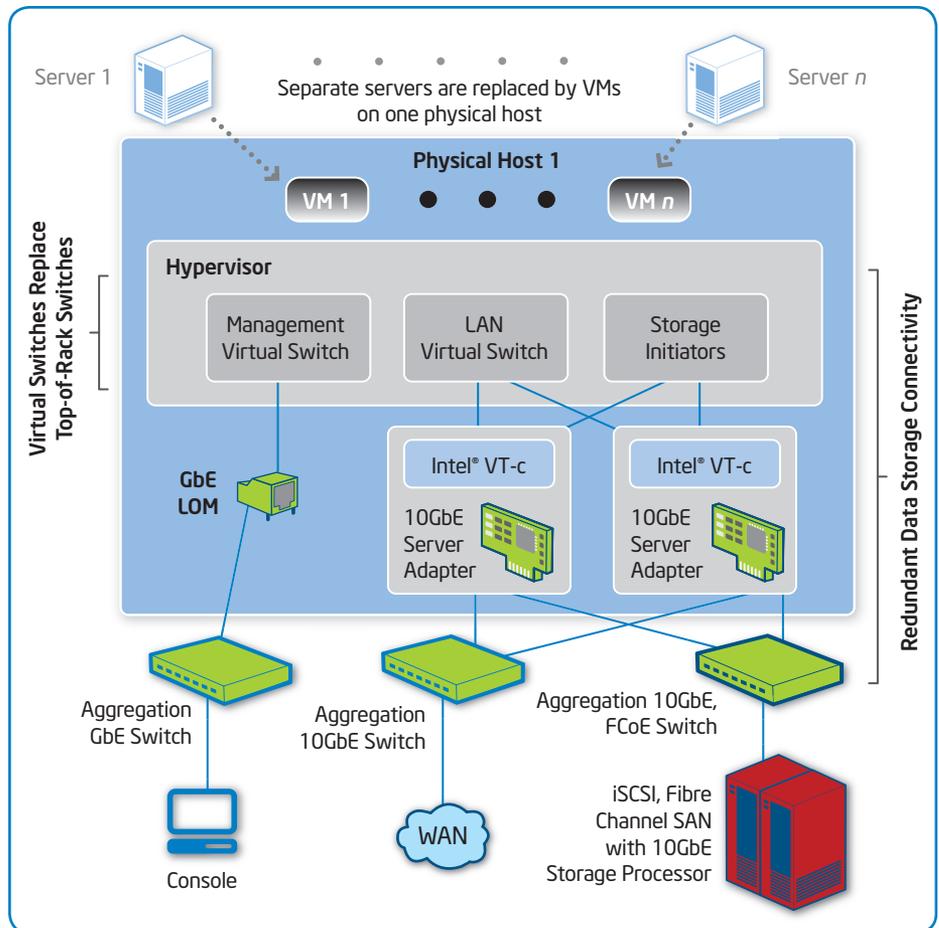


Figure 3. Intel® Virtualization for Connectivity offloads switching and sorting from the processor to the network silicon, for optimized performance.

sorting incoming network I/O data, as well as improving fairness among VMs. This capability provides multiple network queues and a hardware-based sorter/classifier built into the network controller, providing optimized I/O performance as well as add-on capabilities such as a firewall.

Intel VT-c enables a single Ethernet port to appear as multiple adapters to VMs. Each VM's device buffer is assigned a transmit/receive queue pair in the Intel Ethernet controller, and this pairing between VM and network hardware helps avoid needless packet copies and route lookups in the hypervisor's virtual switch. The result is less data for the hypervisor to process and an overall improvement in I/O performance.

Intel VT-c includes Virtual Machine Device Queues (VMDq)² and Single Root I/O Virtualization (SR-IOV) support (please refer to the sidebars in this document for details on these technologies). Both enable dynamic allocation of bandwidth for optimal efficiency in various traffic conditions and are engineered specifically to provide optimal flexibility in migration as well as elasticity of bandwidth.

Proprietary hardware, on the other hand, takes the approach of splitting up the 10GbE bandwidth by advertising separate functions for virtual connections. This method inherently limits the virtual ports to a small number of connections and fails to address the elasticity needs of the network by prescribing a static bandwidth allocation per port.

I/O VIRTUALIZATION WITH INTEL® VT-c

Intel® Virtualization Technology for Connectivity (Intel® VT-c) supports both approaches to I/O virtualization that are used by leading hypervisors, as shown in Figure B:

- **Virtual Machine Device Queues (VMDq)** works in conjunction with the hypervisor to take advantage of the on-controller sorting and traffic steering of the Receive and Transmit queues to provide balanced bandwidth allocation, increased I/O performance, and support for differential traffic loads among VMs.
- **Single Root I/O Virtualization (SR-IOV)** is a PCI-SIG standard that virtualizes physical I/O ports of a network controller into multiple virtual functions (VFs), which it maps to individual VMs, enabling them to achieve near-native network I/O performance by making use of direct access.

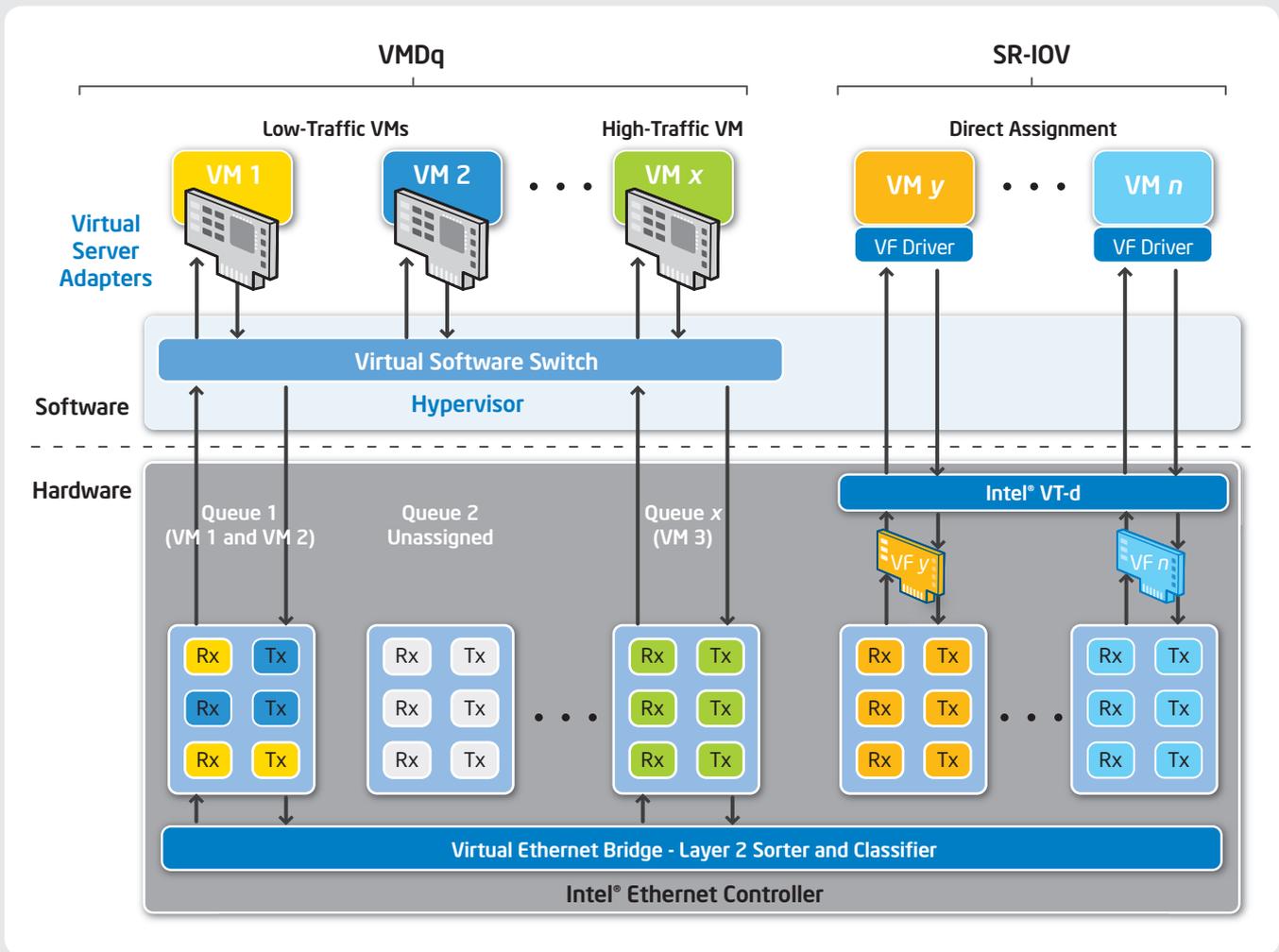


Figure B. Intel® VT-c supports I/O virtualization with both VMDq and SR-IOV, both of which enhance network performance dramatically for cloud environments.

Best Practice #4: Enable a Solution That Works with Multiple Hypervisors

As is typical in IT organizations, requirements and solutions evolve and change over time. Supporting multiple hypervisors in a private cloud is a key enabling capability for optimizing and transitioning your data center, as shown in Figure 4.

This need arises when different groups in a company have implemented different technologies or policies, driven by factors that include preference and licensing costs. Different hypervisors take different approaches to I/O virtualization—Intel Ethernet provides a single solution by

means of Intel VT-c, providing both VMDq and SR-IOV support to be used where appropriate:

- **VMware NetQueue* and Microsoft VMQ** are API implementations of the VMDq component of Intel VT-c. VMDq is a virtualized I/O technology built into all Intel Ethernet 10Gb Server Adapters. The performance of Intel Ethernet controllers enabled with VMDq compares favorably to competing solutions.
- **XenServer* and Kernel Virtual Machine (KVM, the virtualization technology used, for example, in Red Hat Enterprise Linux)** use the SR-IOV support component of

Intel VT-c. SR-IOV delivers performance benefit by completely bypassing the hypervisor and utilizing the bridge built into the server adapter.³

Because Intel VT-c supports both these I/O virtualization technologies, IT organizations can standardize on Intel Ethernet 10 Gigabit Server Adapters for all their physical hosts, knowing that I/O is optimized regardless of which hypervisor is used. Additionally, if a server is reprovisioned with a different hypervisor at some point, no changes to the server adapter are required when utilizing Intel VT-c.

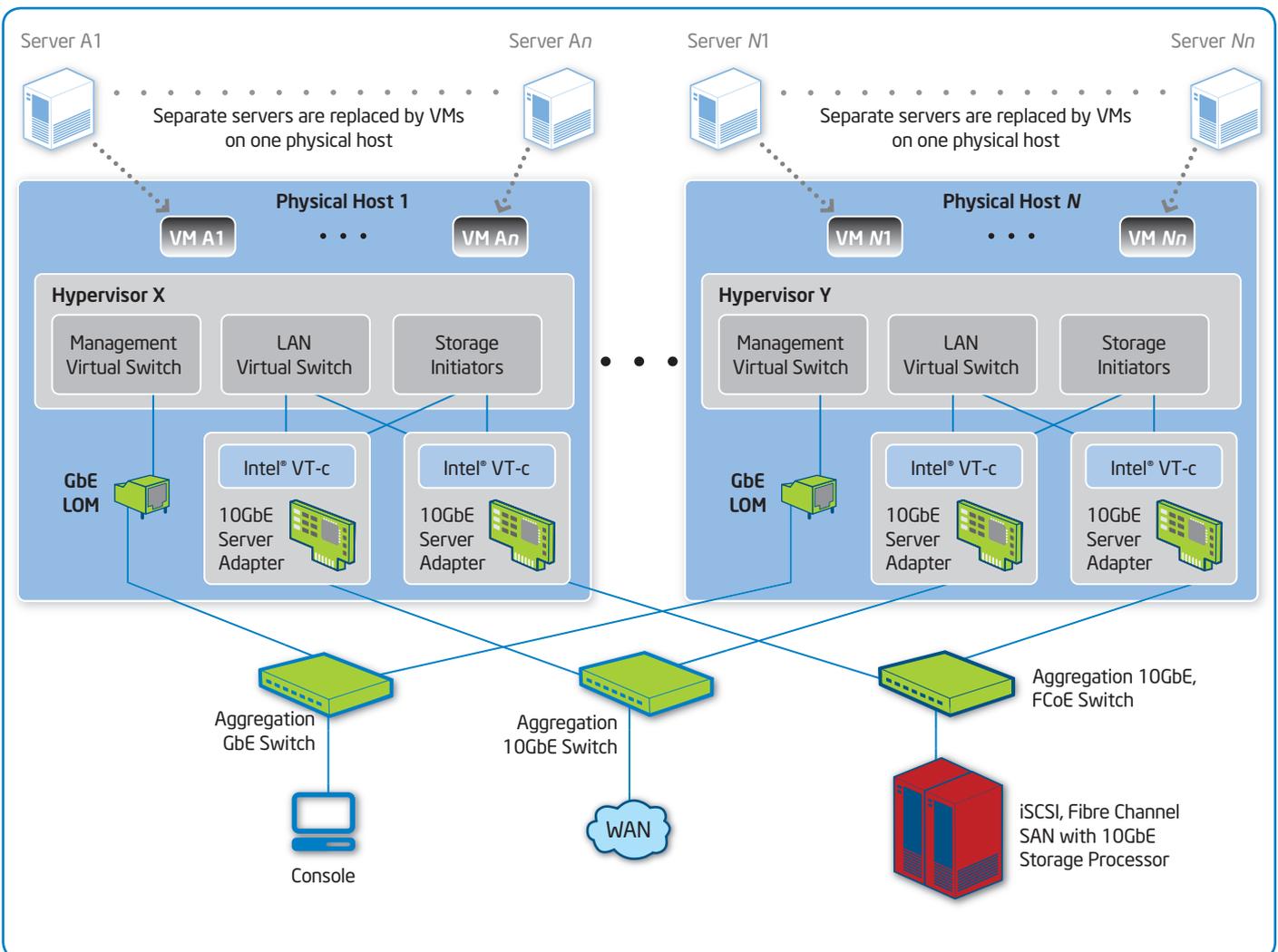


Figure 4. Intel® Virtualization for Connectivity supports combinations of all major hypervisors in the same environment by providing both Virtual Machine Device Queues and Single Root I/O Virtualization support.

Best Practice #5: Utilize QoS for Multi-Tenant Networking

The daily reality for IT organizations is that they must accommodate the needs of their customers that conduct business with IT’s data center assets. In the case of a private cloud, the IT customers are business units, such as engineering, sales, marketing, and finance, within the company. Those customers often have different QoS requirements, represented in Figure 5 as bronze (low), silver (medium), and gold (high).

QoS can be divided into two areas: “on the wire” and “at the guest level.” In this context, “on the wire” refers

to QoS considerations related to IT’s responsibility to guarantee virtual network services such as storage, management, and applications. This includes the management of multiple traffic classes that exist on a single physical port and that have individual requirements in terms of latency, priority, and packet delivery.

DCB is the set of IEEE open standards that define how these traffic classes are implemented and enforced on a single wire:

- Data Center Bridging eXchange (DCBX)
- Enhanced Transmission Selection (ETS)
- Priority Flow Control (PFC)
- Quantized Congestion Notification (QCN)

The use of DCB is preferable as a means of allocating guaranteed amounts of bandwidth to specific traffic classes, as opposed to multiple physical ports or statically assigning bandwidth. Solutions that statically assign bandwidth typically do not allow the bandwidth to exceed the set limit, even if there is available bandwidth on the wire. DCB provides an advantage here because it limits the maximum bandwidth only in cases of contention across traffic classes.

The “guest level” aspect of QoS involves different levels of service for customers and groups, largely by means of bandwidth limiting and services within the

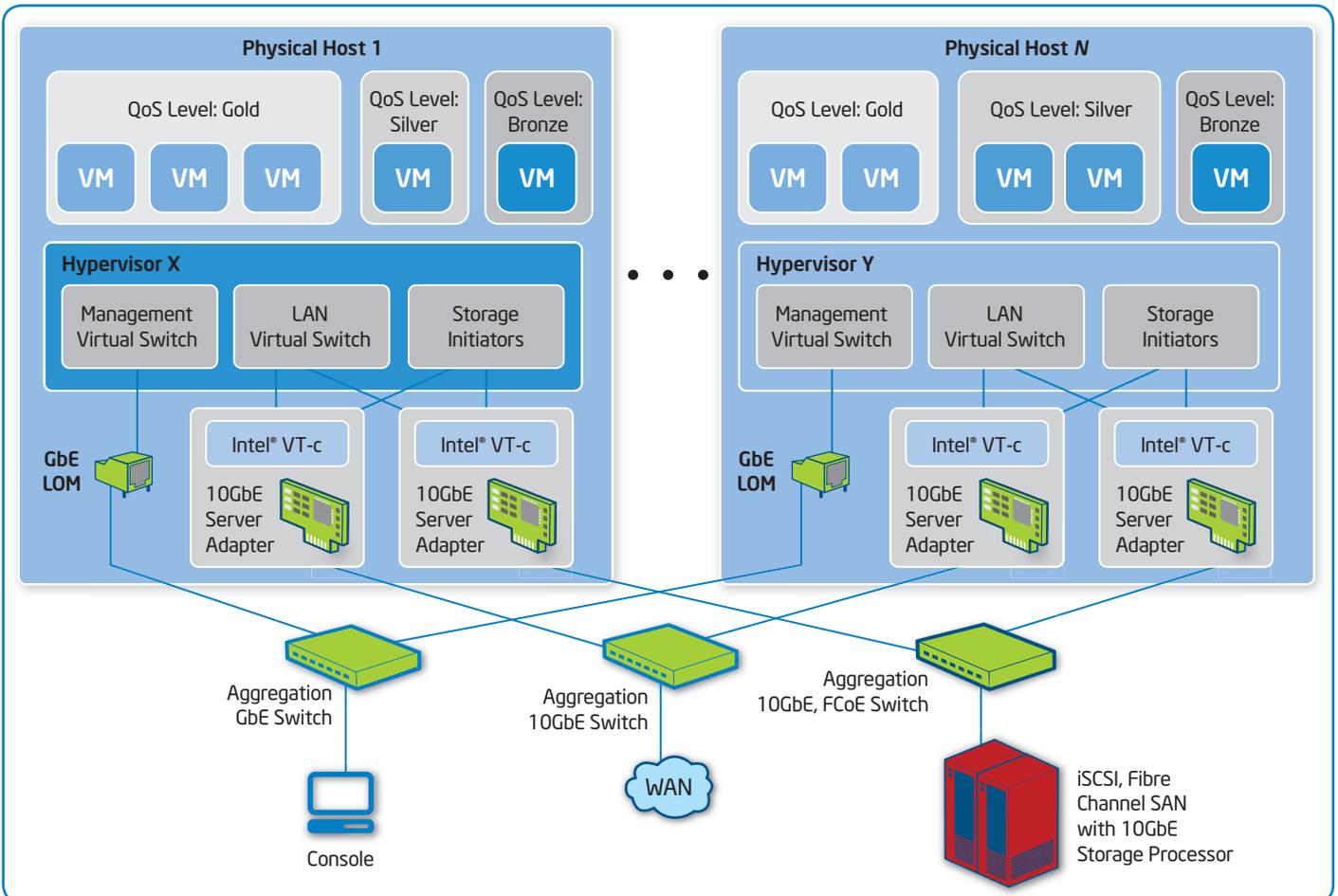


Figure 5. Network policy, traffic shaping, and scheduling enable robust quality-of-service support.

hypervisor. While 10GbE typically provides ample bandwidth for these multiple traffic streams to coexist, policies must be put in place to ensure appropriate allocation of network resources in the event of traffic spikes and congestion.

Bandwidth limiting within a VM or virtual switch enables separate guarantees to be set up on a per-VM (per-customer) basis. They can be controlled either in software or in hardware. VMware and Microsoft Hyper-V* enable I/O limiting from the hypervisor, whereas Xen* and KVM enable I/O limiting from the hardware/driver. Intel Ethernet supports hardware-based rate

limiting in the network controller on a per-VM basis. This flexibility adds further to the ability of the IT organization to tailor services to the needs of individual customer groups.

QoS support with Intel Ethernet 10Gb Server Adapters works by means of virtual network adapters. Intel Ethernet controllers dynamically assign bandwidth to meet QoS requirements using the following features and capabilities:

- **Receive(Rx)/Transmit(Tx) round-robin scheduling** assigns time slices to each VF in equal portions and in a circular fashion for Rx/Tx operations.

- **Traffic steering** enables the virtual Ethernet bridge to sort and classify traffic into VFs or queues.

- **Transmit-rate limiting** throttles the rate at which a VF can transmit data.

- **Traffic isolation** assigns each process or VM a VF with VLAN support.

- **On-chip VM-VM switching** enables the Virtual Ethernet Bridge to conduct PCI Express speed switching between VMs.

Conclusion

To fully realize the agility and cost benefits associated with cloud computing, IT departments must incorporate into their architectures best practices such as those described in this paper:

Best Practice #1: Consolidate Ports and Cables Where Appropriate. Consolidating multiple GbE server adapters onto 10GbE further simplifies the environment, for easier management and lower risk of device failure. The increased bandwidth available from 10GbE also improves handling of workload peaks.

Best Practice #2: Converge Data and Storage onto Ethernet Fabrics. Using iSCSI or FCoE on Intel Ethernet Server Adapters instead of native Fibre Channel to pass storage traffic reduces the capital expense associated with buying host bus adapters and the higher per-port cost of Fibre Channel switching. It also reduces power requirements and simplifies the overall environment by reducing two separate fabrics to just one.

Best Practice #3: Maximize I/O Virtualization Performance and Flexibility. Using Intel VT-c capabilities, including VMDq and SR-IOV support, offloads packet switching and sorting from the processor to the network silicon, making processor resources available for other work.

Best Practice #4: Enable a Solution That Works with Multiple Hypervisors. Because Intel Ethernet Server Adapters support both VMDq and SR-IOV technologies, they can provide the functionality required by various hypervisors, for optimal hardware flexibility.

Best Practice #5: Utilize QoS for Multi-Tenant Networking. Features built into Intel Ethernet Server Adapters also provide excellent support for QoS, dynamically assigning bandwidth as needed to various traffic classes. This functionality is a key component of being able to provide needed service levels to multiple clients in a multi-tenant environment.

Best practices to implement an optimized, reliable private cloud network using Intel Ethernet 10 Gigabit Server Adapters enable IT departments to better align their strategies with those of the business as a whole. As a result, IT is more efficient at building success for its customer organizations.

Take the Next Step

For more information about Intel® Ethernet 10GbE Server Adapters, visit
www.intel.com/go/10GbE

To learn more about Intel's efforts to make cloud infrastructure
simple, efficient, and secure, visit
www.intel.com/go/cloud

¹ Results provided by EMC Corp and Yahoo Inc.

² Available on select Intel® Ethernet Controllers; see http://www.intel.com/network/connectivity/vtc_vmdq.htm.

³ Test Configuration: Ixia IxChariot® v7.1; 16 clients per port under test; high-performance throughput script; file size = 64 to 1 K; 1,000,000 / 2K+; 10,000,000 bytes; buffer sizes = 64 bytes to 64 KB; data type – zeroes; data verification disabled; nagles disabled

System Under Test: Intel® Server Board S5520HC (formerly codenamed Hanlan Creek); two Intel® Xeon® processors X5680 (12 M Cache, 3.33 GHz, 6.40 GT/s Intel® QuickPath Interconnect); Intel® 5520 Chipset (formerly codenamed Tylersburg); RAM: 12 GB DDR3 @ 1333 MHz; BIOS: 0050; Windows Server® 2008 R2 x64.

Clients: SuperMicro 6015T-TV; two dual-core Intel® Xeon® processors 5160 @ 3.0 GHz; 2 GB RAM; Intel® PRO/1000 PT Dual Port Server Adapter - v9.12.13.0 driver; Windows Server® 2003 SP2 x64.

Network Configuration: Force10 Networks ExaScale® E1200i switch; clients connected @ 1 Gbps.

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. UNLESS OTHERWISE AGREED IN WRITING BY INTEL, THE INTEL PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE INTEL PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined." Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request. Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order. Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or by visiting Intel's Web site at www.intel.com.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to <http://www.intel.com/performance>.

Copyright © 2011 Intel Corporation. All rights reserved. Intel, the Intel logo, and Xeon are trademarks of Intel Corporation in the U.S. and other countries.

*Other names and brands may be claimed as the property of others.

Printed in USA

0811/BY/MESH/PDF

 Please Recycle

325931-001US

