

October 15-17, 2013

COMPUTERWORLD SNIA[®]
SNW

Intelligent Architecture for
the Data-Driven Business

I/O Virtualization

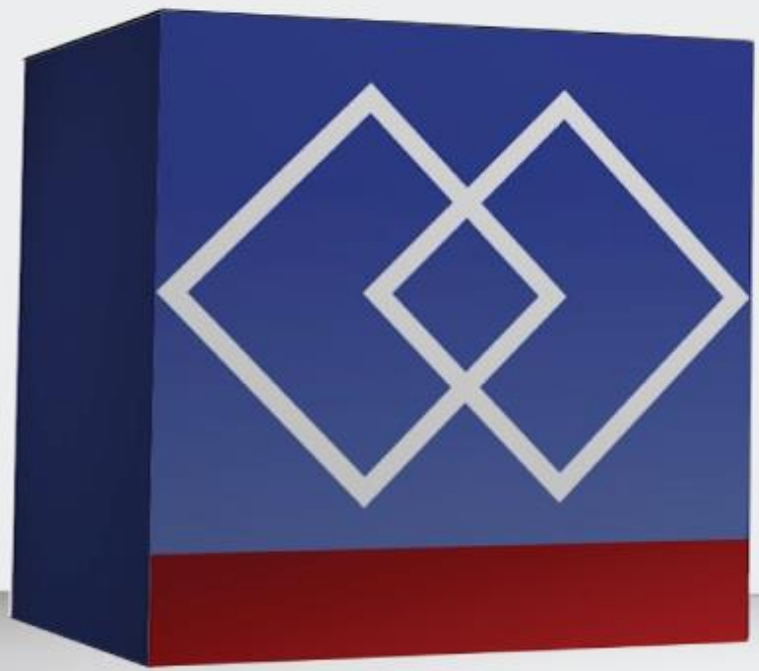
The Next Virtualization Frontier

Dennis Martin, President, Demartek



Agenda

- ◆ **Demartek – About Us**
- ◆ **Why Do We Virtualize?**
- ◆ **I/O Virtualization – What Is It?**
- ◆ **Virtualizing the PCIe Bus**
- ◆ **SR-IOV, MR-IOV and Hybrid**
- ◆ **NPIV and Virtual Fibre Channel**
- ◆ **SR-IOV Performance Results**
- ◆ **References**



About Demartek

**Demartek Releases
Video Detailing Services Provided.**

Curious about what we do? Here's your answer.

 **Demartek**

SERVICES
[Explore >>](#)

[View the Demartek Video Library](#)

Why Do We Virtualize?

- ◆ **De-couple the logical from the physical**
 - **Hardware can be split into smaller logical units**
 - **Hardware can be represented as multiple units**
 - **Hardware can be combined into larger units**
- ◆ **Want to use computing resources more effectively, especially the under-utilized assets**
- ◆ **Improves deployment time**
- ◆ **Allows expensive resources to be shared, or shared more widely**

Examples Available Today

◆ NIC Teaming or Bonding

- Logically join two or more NIC ports and present them as one NIC port
- Can be used for failover or load balancing
- Often, but not always, specific to one brand of NIC

◆ NPIV

- FC HBA that can present multiple virtual ports using the same physical port
- Can be used in VM environments

I/O Virtualization – What Is It?

- ◆ Virtualizing the *I/O path* between a server and an external device
- ◆ Can apply to anything that uses an adapter in a server, such as:
 - Ethernet Network Interface Cards (NICs)
 - Disk Controllers (including RAID controllers)
 - Fibre Channel Host Bus Adapters (HBAs)
 - Graphics/Video cards or co-processors
 - SSDs mounted on internal cards

Virtualizing the PCIe Bus

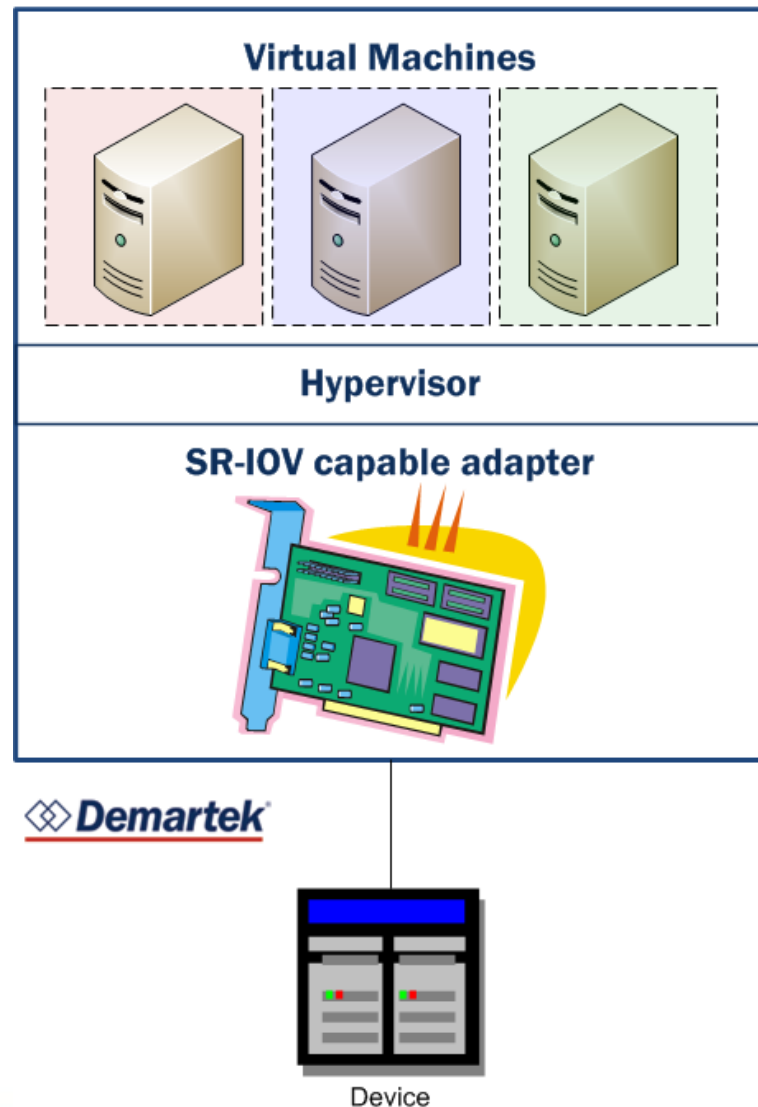
- ◆ **In June 2008, the PCI-SIG®⁷, the Special Interest Group responsible for PCI Express® (PCIe®) industry-standard I/O technology, announced the completion of the PCI-SIG I/O Virtualization (IOV) suite of specifications**
- ◆ **Works with system virtualization technologies**
- ◆ **Allows multiple operating systems to natively share PCI-Express devices**

Virtualizing the I/O Path

Single-Root (SR-IOV)

- ◆ Multiple VMs sharing one I/O adapter
- ◆ Bandwidth of the I/O adapter is shared among the VMs
- ◆ Virtual adapters created and managed by SR-IOV adapter (not hypervisor)
- ◆ Improved performance for VMs and their apps (near-native)

SR-IOV Architecture



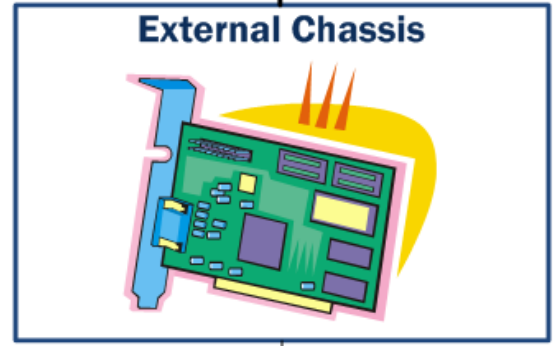
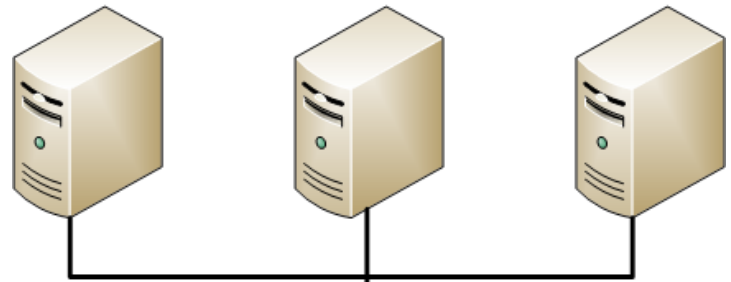
Virtualizing the I/O Path

Multi-Root (MR-IOV)

- ◆ Multiple servers & VMs sharing one I/O adapter
- ◆ Bandwidth of the I/O adapter is shared among the servers
- ◆ The I/O adapter is placed into a separate chassis
- ◆ Bus extender cards are placed into the servers

MR-IOV Architecture

Physical Servers



 **Demartek**



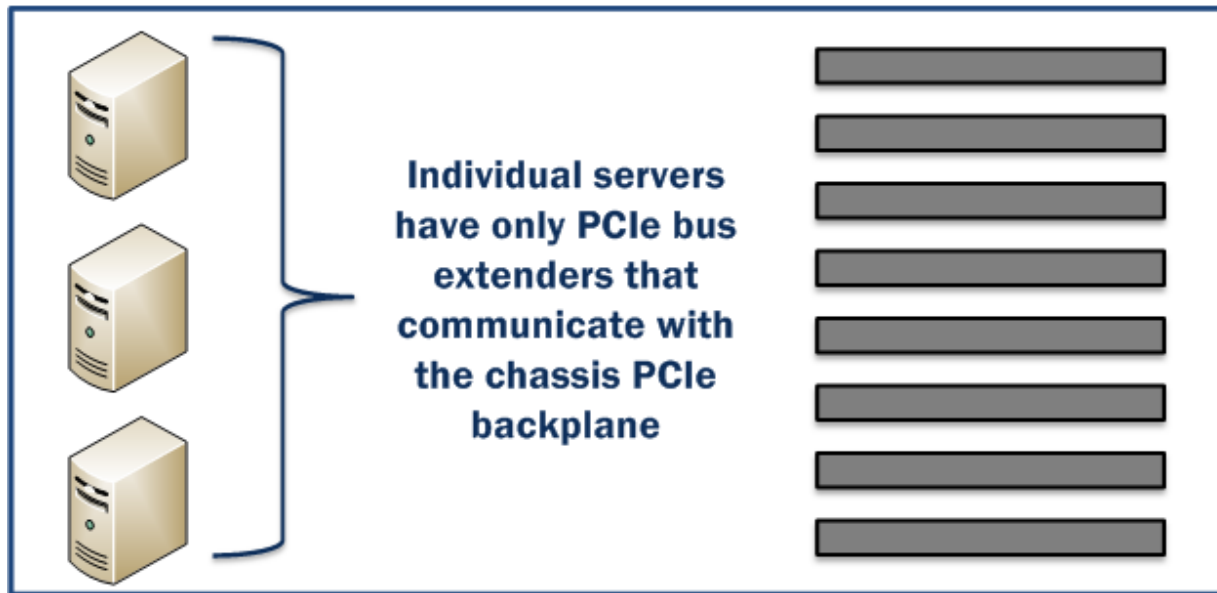
Device

Virtualizing the I/O Path

Hybrid SR-IOV/MR-IOV

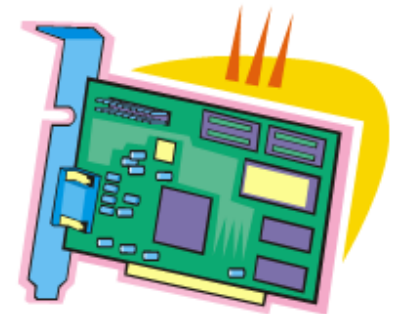
- ◆ **Common PCIe backplane for several servers in new hybrid blade/rack server**

Hybrid SR-IOV/MR-IOV Architecture



Hybrid Blade Server Chassis

PCIe busses can be assigned to any one server



Uses standard PCIe adapters including SR-IOV capable

Hairpin Turns

- ◆ **In an IOV-capable environment, traffic can be sent out of one virtual adapter and received into another virtual adapter**
- ◆ **These two virtual adapters could reside on the same physical adapter, resulting in a “hairpin turn”**
- ◆ **The IOV adapters or IOV switches could act as LAN or SAN switches within the PCIe fabric (at lower cost)**

Virtual Functions (VF)

- ◆ **Virtual functions are the way that the adapter makes multiple versions of itself visible to the Hypervisor**
- ◆ **The Hypervisor assigns a VF to a guest**
 - Many VFs can be created per physical port
 - Hypervisor has no visibility into the VF
- ◆ **Guest sees a new adapter that it can use for anything that adapter can do**

SR-IOV Today

- ◆ **SR-IOV adapter availability**
 - **Ethernet NICs (1Gb and 10Gb)**
 - **FC HBAs**
 - **Some RAID controller prototypes**
- ◆ **Hypervisor support (Ethernet only)**
- ◆ **Citrix XenServer**
 - **RHEL 6 KVM**
 - **VMware 5.1 and higher**
 - **vMotion is not supported with SR-IOV**
 - **Windows Server 2012 (Hyper-V and guest)**
 - **Live Migration is supported with SR-IOV**

SR-IOV Hardware Dependencies

◆ Hardware Checklist

- Processor (CPU) support
- Motherboard support (chipset, etc.)
- BIOS support
- SR-IOV capable NIC

◆ Many, but not all PCIe 2.0 and 3.0 servers will meet all of these criteria

SR-IOV Software Dependencies

- ◆ Hypervisor IOV-enabled virtual switch
- ◆ VF driver for the guest O.S.
- ◆ Windows registry: `lovEnableOverride`
- ◆ Set in the parent (Hyper-V) partition
- ◆ Windows is more strict about making sure that the hardware checklist is satisfied

SR-IOV Bottom Line

- ◆ **High traffic VMs can be assigned directly to a PCI Express Virtual Function provided by SR-IOV support in the OS, Server Platform and Network Controller by-passing the Hypervisor and Virtual Switch**
- ◆ **Advantage: Guest sees higher performance than it would without SR-IOV**
- ◆ **Disadvantage: Hypervisor has no visibility**

NPIV & Virtual Fibre Channel

- ◆ **N-Port ID Virtualization (NPIV) for Fibre Channel behaves in a similar manner to SR-IOV for Ethernet**
- ◆ **When enabled, the FC HBA has VFs that can be assigned to guests**
- ◆ **Guest VM sees a virtual FC HBA and has access to same FC infrastructure as a physical server**

Virtual Fibre Channel – Hyper-V

The screenshot shows a window titled "Virtual SAN Manager for DMRTK-SRVR-K". On the left is a navigation pane with the following items:

- Virtual Fibre Channel SANs
 - New Fibre Channel SAN
- Global Fibre Channel Settings
 - World Wide Names
 - C003FFDABA5F0000 to C003FFDA...

The main area displays a wizard titled "Create Virtual Fibre Channel Storage Area Network". It includes the following text and elements:

- Icon of a lightbulb.
- Text: "Create Virtual Fibre Channel Storage Area Network"
- Text: "Click Create to add a virtual Fibre Channel storage area network (SAN)."
- A text input field containing "Virtual Fibre Channel SAN".
- A "Create" button.
- Text: "A virtual Fibre Channel SAN groups physical HBA ports together. You can add a virtual Fibre Channel adapter to a virtual machine and connect it to a virtual SAN."

Virtual Fibre Channel – Hyper-V

Virtual SAN Manager for DMRTK-SRVR-K

Virtual Fibre Channel SANs

- New Fibre Channel SAN
- FC-1**

Global Fibre Channel Settings

- World Wide Names
C003FFDABA5F0000 to C003FFDA...

New Fibre Channel SAN

Name: FC-1

Notes:

	WWNN	WWPN	Status
<input checked="" type="checkbox"/>	20008C7CFF0B8E81	10008C7CFF0B8E81	'New Fibre Channel SAN'
<input type="checkbox"/>	20008C7CFF0B8E80	10008C7CFF0B8E80	Available
<input type="checkbox"/>	200000051E0F8521	100000051E0F8521	Available
<input type="checkbox"/>	200000051E0F8520	100000051E0F8520	Available

Remove virtual SAN

Virtual Fibre Channel – Guest VM

Settings for Win Server 2012 on DMRTK-SRVR-K

Win Server 2012

- Hardware
 - Add Hardware
 - BIOS
 - Boot from CD
 - Memory
 - 512 MB
 - Processor
 - 1 Virtual processor
 - IDE Controller 0
 - Hard Drive
 - HD1.vhdx
 - IDE Controller 1
 - DVD Drive
 - None
 - SCSI Controller
 - Network Adapter
 - Not connected
 - COM 1
 - None
 - COM 2
 - None
 - Diskette Drive
 - None
- Management

Add Hardware

You can use this setting to add devices to your virtual machine. Select the devices you want to add and click the Add button.

- SCSI Controller
- Network Adapter
- Legacy Network Adapter
- Fibre Channel Adapter**
- RemoteFX 3D Video Adapter

Add

You can use a Fibre Channel adapter to access Fibre Channel based storage directly from the guest operating system. Integration services are required in the guest operating system to this access. Do not attach a system disk to a Fibre Channel adapter. System disks must be attached to an IDE controller.

Virtual Fibre Channel – Guest VM

The screenshot shows the Hyper-V settings for a virtual machine named 'Win Server 2012' on host 'DMRTK-SRVR-K'. The left-hand navigation pane is expanded to 'Hardware', and 'Fibre Channel Adapter FC-1' is selected. The main pane displays the configuration for the 'Fibre Channel Adapter'. It includes a 'Virtual SAN' dropdown menu set to 'FC-1', an 'Edit Addresses' button, and a 'Port addresses' section with two address sets (A and B). Each address set has input fields for 'World Wide Node Name (WWNN)' and 'World Wide Port Name (WWPN)'. Address set A has WWNN 'C003FF0000FFFF00' and WWPN 'C003FFDABA5F0000'. Address set B has WWNN 'C003FF0000FFFF00' and WWPN 'C003FFDABA5F0001'. A 'Create Addresses' button is located below the address sets. At the bottom, there is a 'Copy' button and a 'Remove' button.

Settings for Win Server 2012 on DMRTK-SRVR-K

Win Server 2012

Hardware

- Add Hardware
- BIOS
Boot from CD
- Memory
512 MB
- Processor
1 Virtual processor
- IDE Controller 0
 - Hard Drive
HD1.vhdx
- IDE Controller 1
 - DVD Drive
None
- SCSI Controller
- Fibre Channel Adapter
FC-1**
- Network Adapter
Not connected
- COM 1
None
- COM 2
None
- Diskette Drive
None

Management

- Name
Win7
- Integration Services
All services offered

Fibre Channel Adapter

You can review and edit the World Wide Names (WWNs) assigned to the Fibre Channel adapter, and connect the adapter to a virtual storage area network (SAN).

Virtual SAN:
FC-1

Click Edit Addresses to edit the port addresses.

Edit Addresses

Port addresses

Address set A:

World Wide Node Name (WWNN): C003FF0000FFFF00

World Wide Port Name (WWPN): C003FFDABA5F0000

Address set B:

World Wide Node Name (WWNN): C003FF0000FFFF00

World Wide Port Name (WWPN): C003FFDABA5F0001

Create Addresses

Click Copy to copy the addresses to the clipboard.

Copy

To remove the adapter from this virtual machine, click Remove.

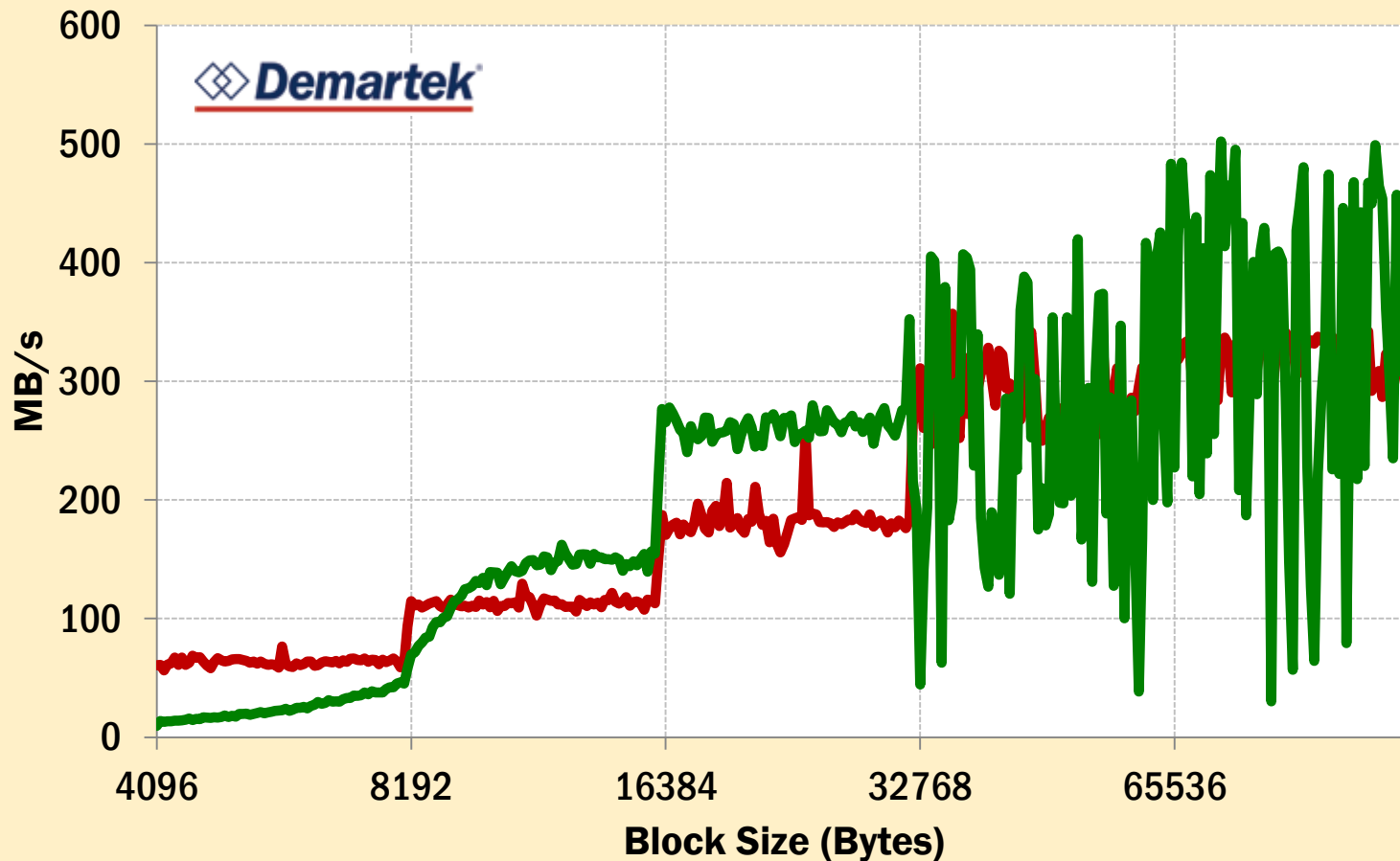
Remove

Performance Examples

SR-IOV Performance Test

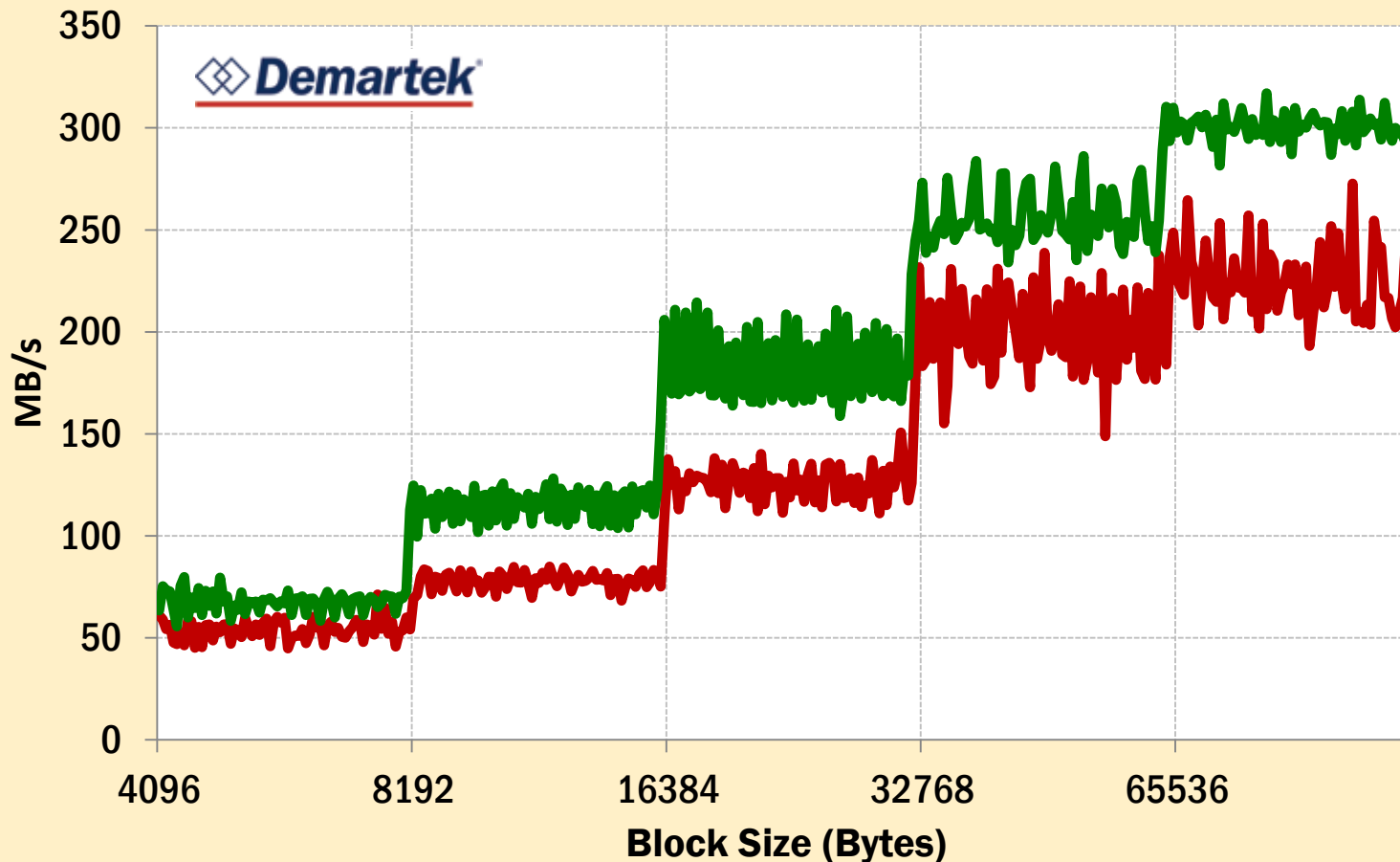
- ◆ **We compared the performance of an SR-IOV capable 10GbE NIC connected via iSCSI to an all-flash storage target**
 - SR-IOV enabled (VF)
 - SR-IOV disabled (shared NIC)
- ◆ **Used vdbench to drive the workload**
- ◆ **Tested with one guest**
- ◆ **Demartek lab server**
 - **2x Intel Xeon X5690, 3.46 GHz, 144 GB RAM with Intel X520 10GbE Server Adapter**

100% Random Read - 10Gb iSCSI



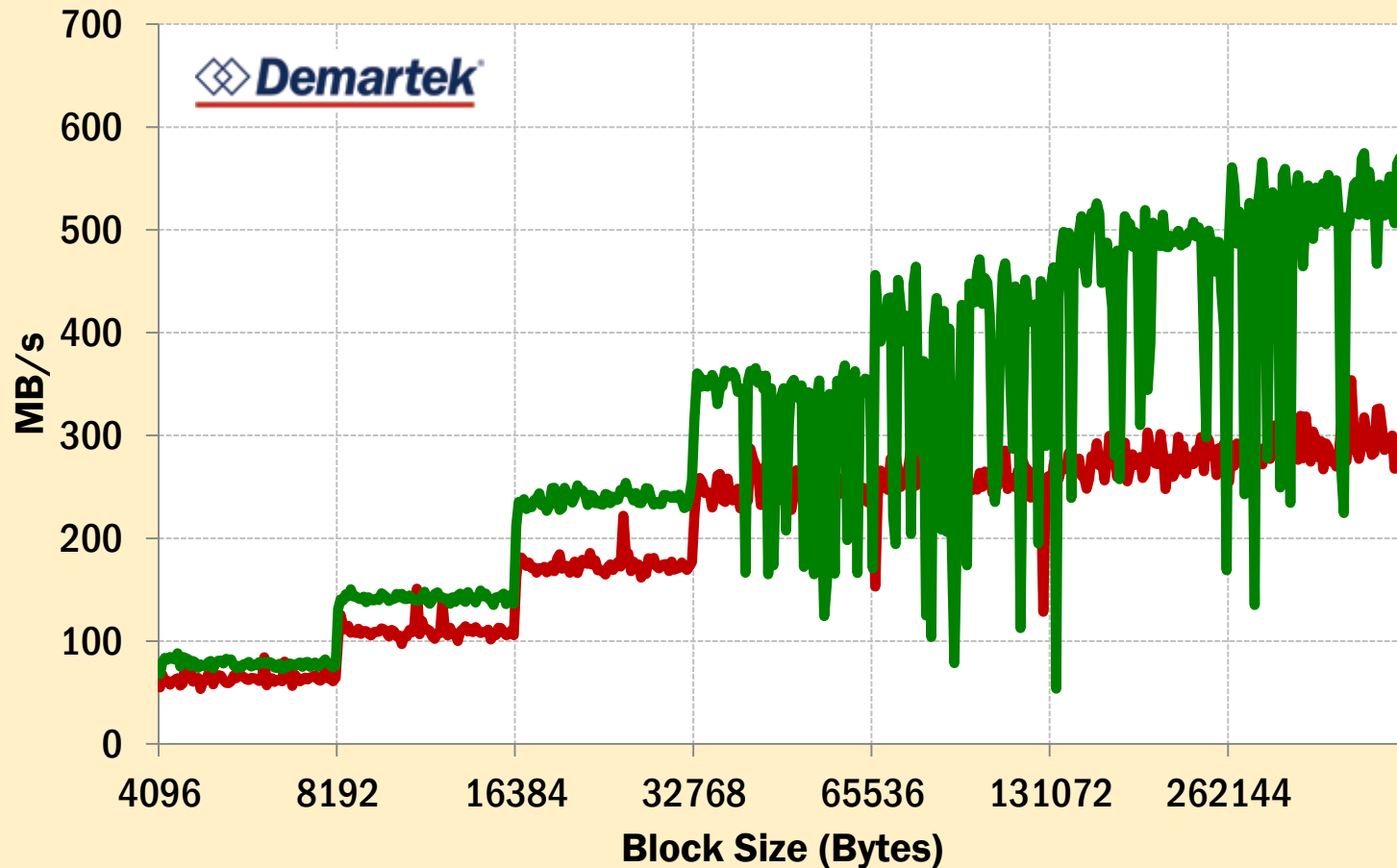
— Shared NIC — NIC Virtual Function

100% Random Write – 10Gb iSCSI



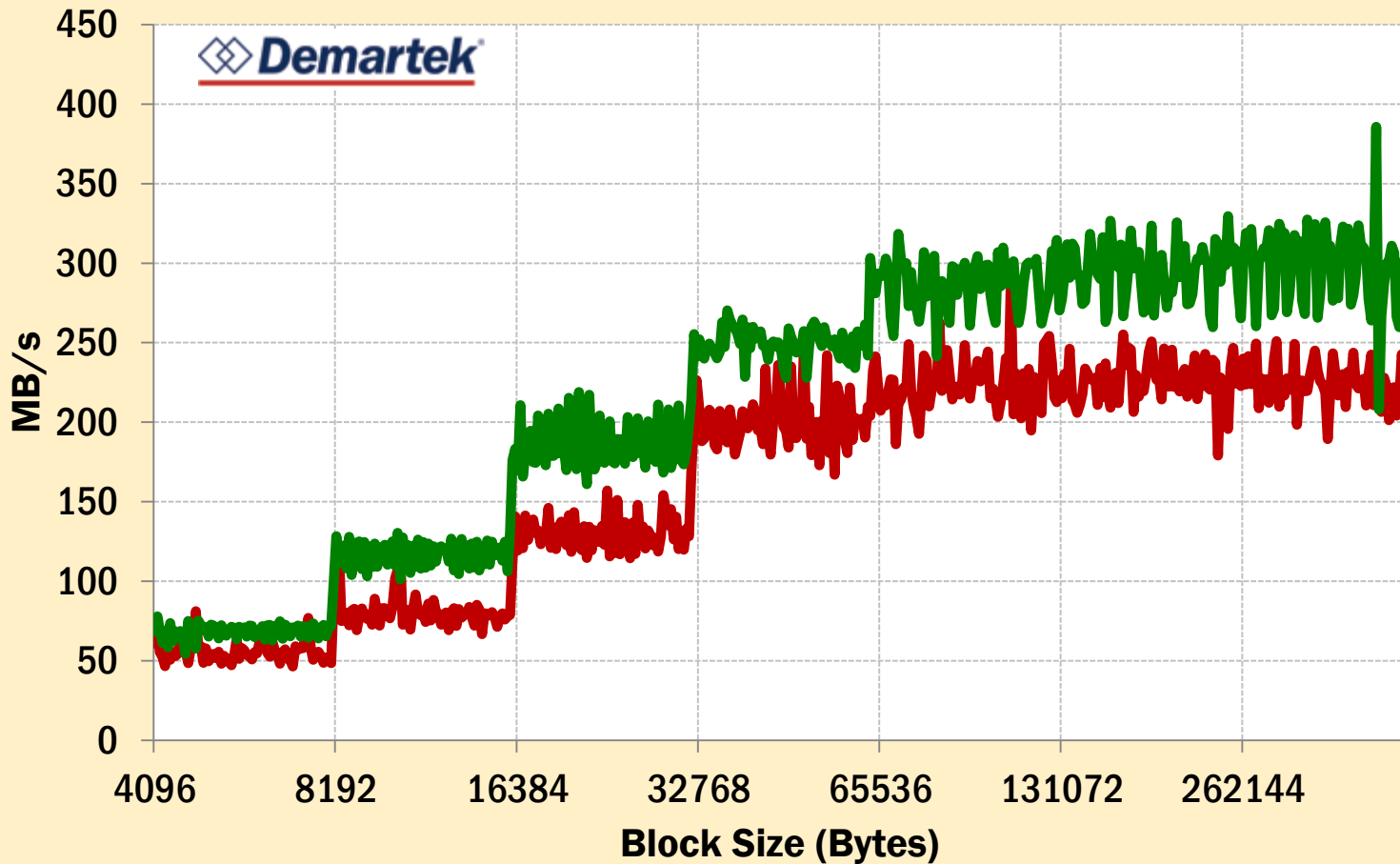
— Shared NIC — NIC Virtual Function

100% Sequential Read – 10Gb iSCSI



— Shared NIC — NIC Virtual Function

100% Sequential Write – 10Gb iSCSI



— Shared NIC — NIC Virtual Function

Demartek References

- ◆ **Demartek Storage Interface Comparison**
 - www.demartek.com/Demartek_Interface_Comparison.html
- ◆ **Demartek SSD Zone**
 - www.demartek.com/SSD
- ◆ **Demartek SSD Deployment Guide**
 - www.demartek.com/Demartek_SSD_Deployment_Guide.html
- ◆ **Demartek Commentary – Horses, Buggies & SSDs**
 - www.demartek.com/Demartek_Horses_Buggies_SSDs_Commentary.html
- ◆ **Demartek Free Monthly Newsletter**
 - www.demartek.com/Newsletter



Thank You!

Dennis Martin, President

dennis@demartek.com

www.linkedin.com/in/dennismartin



(303) 940-7575

www.demartek.com

<http://twitter.com/Demartek>

www.youtube.com/Demartek

Skype: Demartek

To learn more about Demartek:

- ◆ Download the Aurasma App (Android/iPhone)
- ◆ Search and follow “Demartek”
- ◆ View image below with viewfinder.



*also on the back of Dennis' business card

Powered by: 
AURASMA