SNIA

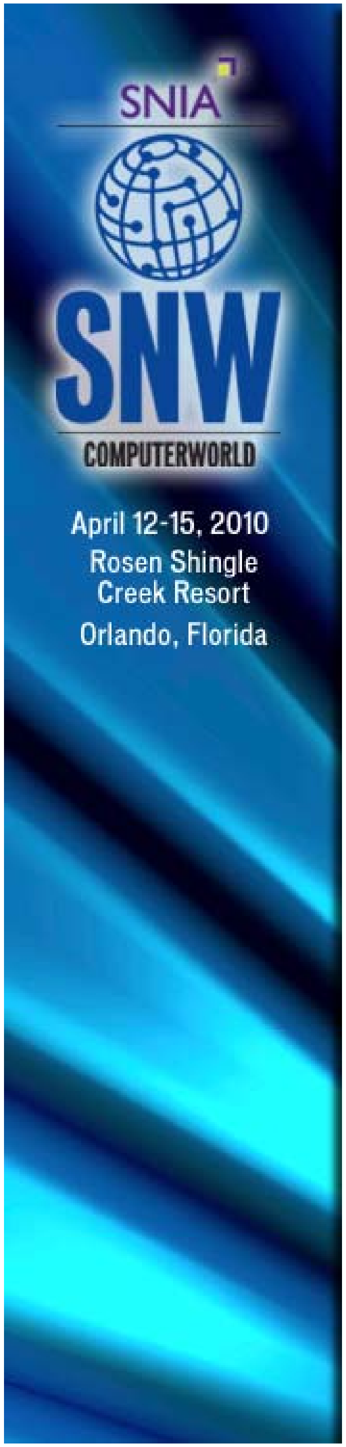
SNW

COMPUTERWORLD

April 12-15, 2010
Rosen Shingle
Creek Resort
Orlando, Florida

I/O Virtualization
The Next Virtualization Frontier

Dennis Martin
President, Demartek



Demartek Company Overview

- Industry analysis with on-site test lab
- Most projects involve use of the lab
- Lab includes servers, networking and storage infrastructure
  - Fibre Channel: 4 & 8 Gbps
  - Ethernet: 1 & 10 Gbps (with FCoE)
  - Servers: at least 8 cores, up to 48GB RAM
  - Virtualization: ESX, Hyper-V, Xen
- Web: www.demartek.com

SNIA
SNW
COMPUTERWORLD
April 12-15, 2010
Rosen Shingle
Creek Resort
Orlando, Florida

Agenda

- Why Do We Virtualize?
- I/O Virtualization – What Is It?
- Virtualizing the PCIe Bus
- Virtualizing the I/O Path (Internal & External)
- Rack Area Networks
- Benefits of I/O Virtualization
- Hairpin Turns
- Discussion of Early Testing

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 of Virtualization

- Virtual Memory
  - Uses memory more effectively
  - Was revolutionary, but now is assumed
- Virtual Storage
  - Presents storage resources in ways not bound to the underlying hardware characteristics
  - Fairly common now
- Virtual Servers
  - Increases typically under-utilized CPU resources
  - Becoming more common

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

Existing Forms of I/O Virtualization

- NIC Teaming
  - A virtual NIC composed of two or more physical NICs
- Virtual LAN
  - Multiple, smaller logical LANs within a physical LAN infrastructure
- Virtual SAN Fabrics
  - Multiple, smaller logical SANs within a physical SAN infrastructure

Virtualizing the PCIe Bus

- In June 2008, the PCI-SIG®, 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

PCIe IOV

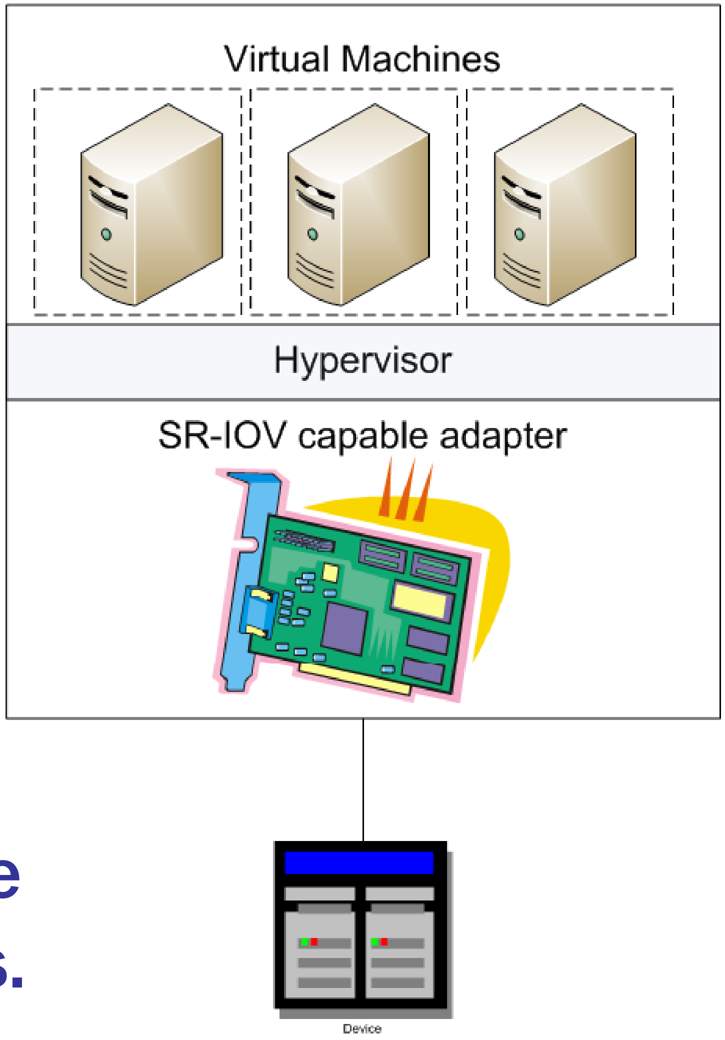
- Address Translation Services (ATS)
  - Enables performance optimizations between an I/O device and the platform's IOMMU
- Single-Root IOV (SR-IOV)
  - Enables multiple guest operating systems to simultaneously access an I/O device without having to trap to the hypervisor on the main data path.
- Multi-Root IOV (MR-IOV)
  - Enables either PCIe or SR-IOV I/O devices to be accessed through a shared PCIe fabric.

Processors and IOV

- Processor vendors:
  - Intel: Virtualization Technology for Directed I/O (VT-d)
  - AMD: Virtualization (AMD-V) Technology
- I/O Virtualization (IOV) includes:
  - I/O device assignment so that I/O devices can be assigned to virtual machines (VMs)
  - DMA remapping for VMs
  - Interrupt remapping for VMs

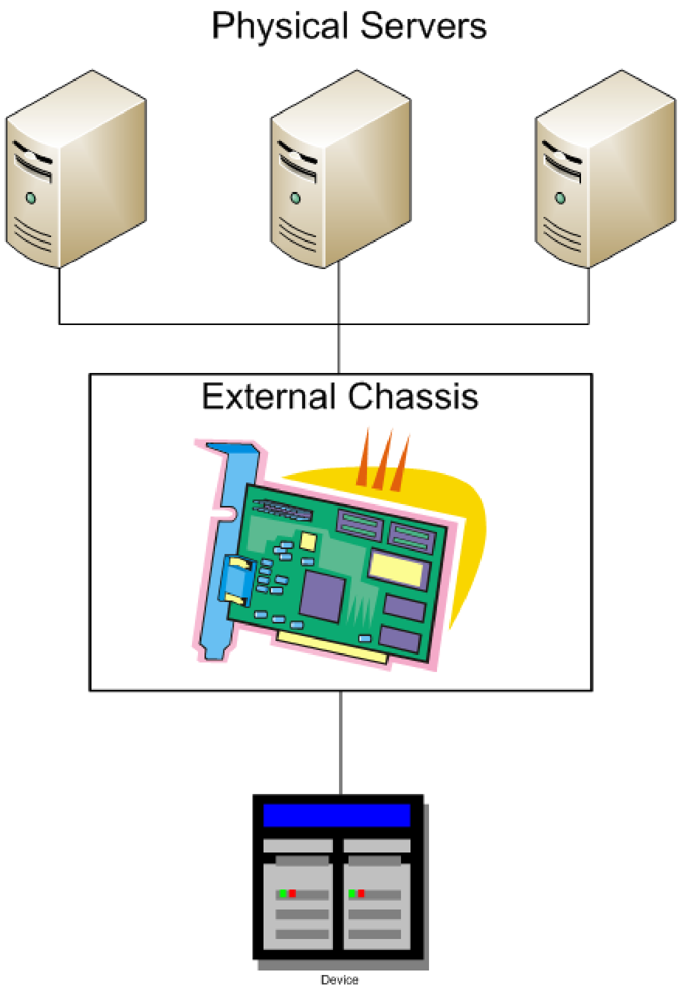
Virtualizing the I/O Path – 1

- 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.

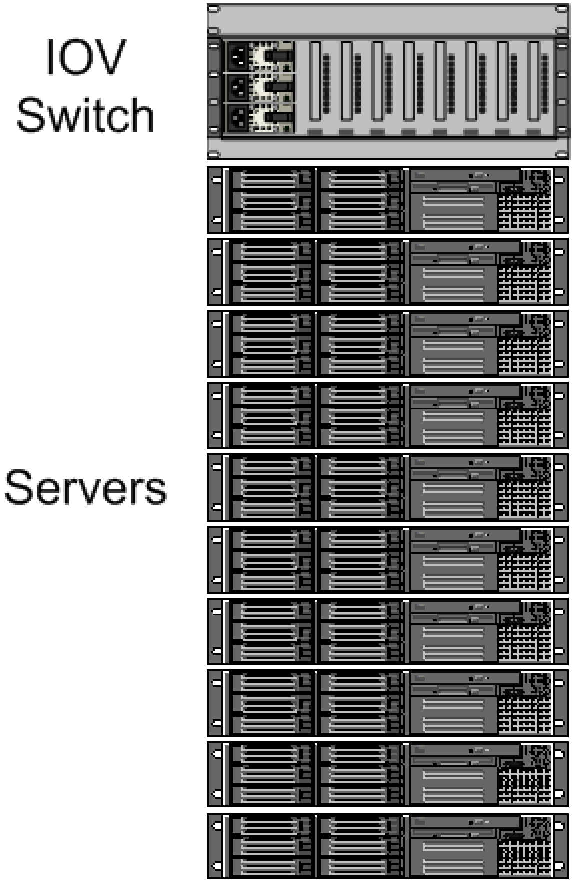


Virtualizing the I/O Path – 2

- 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



Rack Area Networking (RAN)



- IOV switch contains NICs, HBAs, etc
- IOV switch connects to LAN, SAN, DAS in other racks
- PCIe fabric and cables within the rack
- Servers have PCIe extender cards

Rack Area Networking (RAN)

- Uses IOV with multiple physical servers (and their guest VMs)
- Uses a top-of-rack IOV switch that supports various types of PCIe adapters
- Gives rack servers within a rack many of the same benefits as blade servers within a blade chassis
- Allows some rack servers to shrink to 2U, 1U or possibly ½U

SNIA
SNW
COMPUTERWORLD
April 12-15, 2010
Rosen Shingle Creek Resort
Orlando, Florida

External Implementations

- Natively extend the PCI-Express bus to an IOV external chassis or IOV switch

- Encapsulate PCI-Express within Infiniband and extend the Infiniband bus to an IOV external chassis or IOV switch
  - Infiniband runs faster than PCI-Express today, so there is enough bandwidth for this technique

SNIA

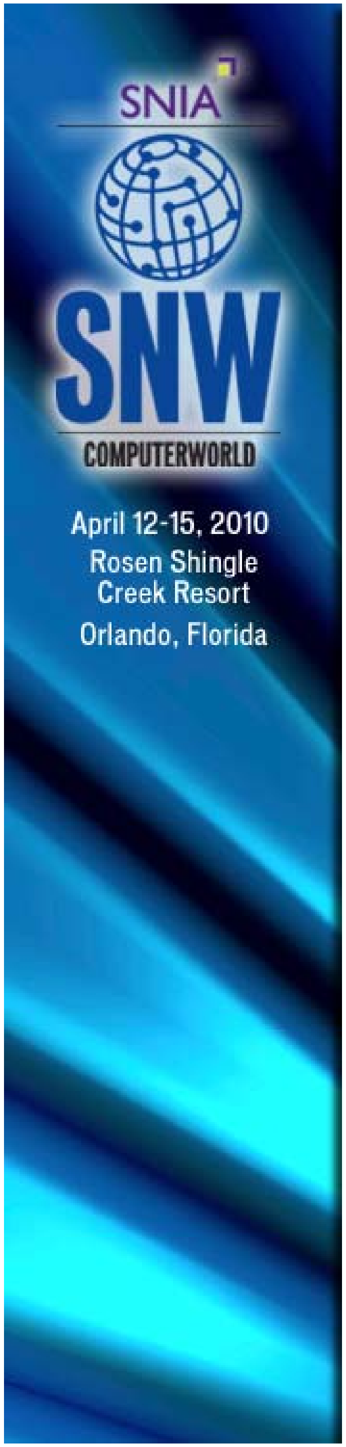
SNW

COMPUTERWORLD

April 12-15, 2010
Rosen Shingle
Creek Resort
Orlando, Florida

Benefits of I/O Virtualization

- Increases utilization of adapters
- Expensive adapters can be shared rather than dedicated to a single server/O.S.
- Decreases power consumption and cooling needs in some cases
- Reduced rack space servers can be deployed in some cases
- O.S. and hypervisor tasks can be offloaded to the adapter, increasing performance
- Some adapter upgrades are external to servers



Move VMs Without a SAN?

- Using IOV, VMs could be connected to their storage via standard disk controllers and DAS.
- Movement of VMs could be done to different physical servers without a traditional SAN

IOV and Other Technologies

- IOV can work with any adapter that uses the PCIe bus and that supports IOV
  - 10GbE NICs
  - FC HBAs
  - FCoE CNAs
  - Disk controllers (with or without RAID)
  - Graphics adapters
  - SSDs mounted on PCIe cards

When and Where to Deploy?

- Expect to see various NICs, HBAs, CNAs and SAS/SATA controllers that support IOV begin to appear in 2010
  - Some prototypes were shown in 2009
- Good candidates are virtual or physical environments that:
  - Can share high-speed adapters
  - Need to share expensive PCIe devices

IOV Management

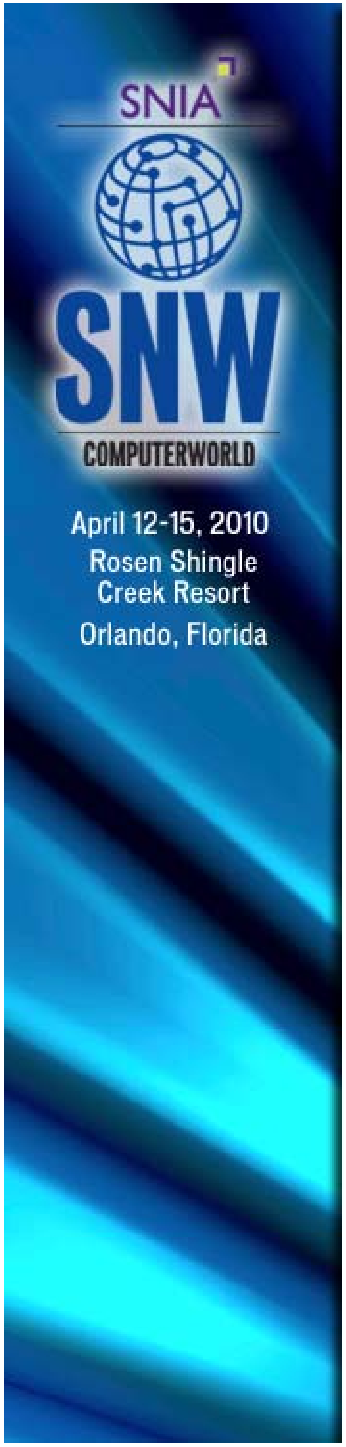
- IOV adapters and paths will no longer be exclusively assigned to individual servers
  - Similar to SAN storage devices that are not owned by an individual server
- O.S. and hypervisor vendors still have work to do to, but are making progress

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)

Testing

- We have begun to test some of these technologies in our on-site lab
- Several smaller companies are building top-of-rack "IOV" switches
- Several larger companies are building IOV-capable adapters
- Look for "I/O Virtualization" or "IOV" on our news page: www.demartek.com/Demartek_news.html



Storage Interface Comparison

- Demartek has compiled a free comparison reference guide of the storage networking interfaces. This reference guide is updated periodically.

www.demartek.com/Demartek_Interface_Comparison.html

SNIA

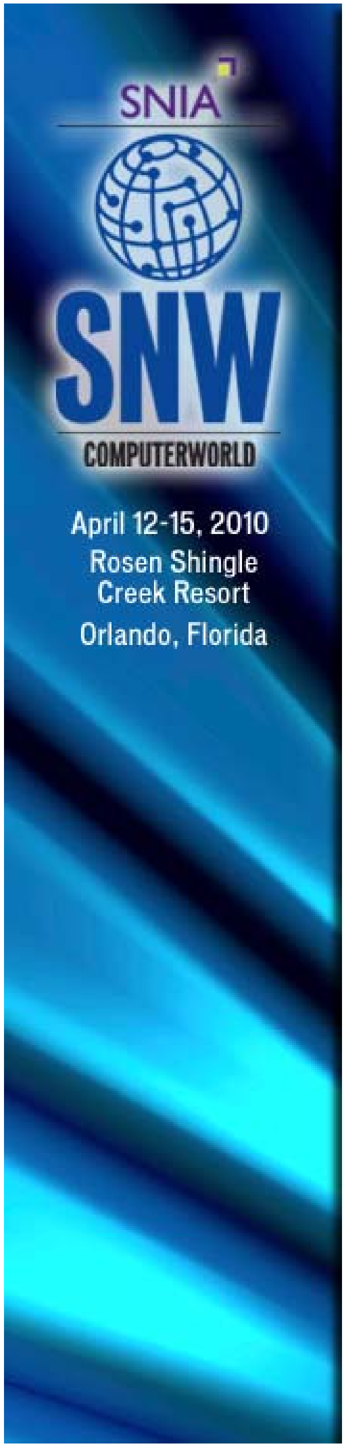
SNW

COMPUTERWORLD

April 12-15, 2010
Rosen Shingle
Creek Resort
Orlando, Florida

Free Monthly Newsletter

- Demartek publishes a free monthly newsletter highlighting recent reports, articles and commentary. Look for the newsletter sign-up at www.demartek.com.



Contact Information

Dennis Martin, President
Demartek
(303) 940-7575
dennis@demartek.com

www.linkedin.com/in/dennismartin
http://twitter.com/demartek