The Future of Virtualization Technology

Stephen Alan Herrod
VP of Technology
VMware

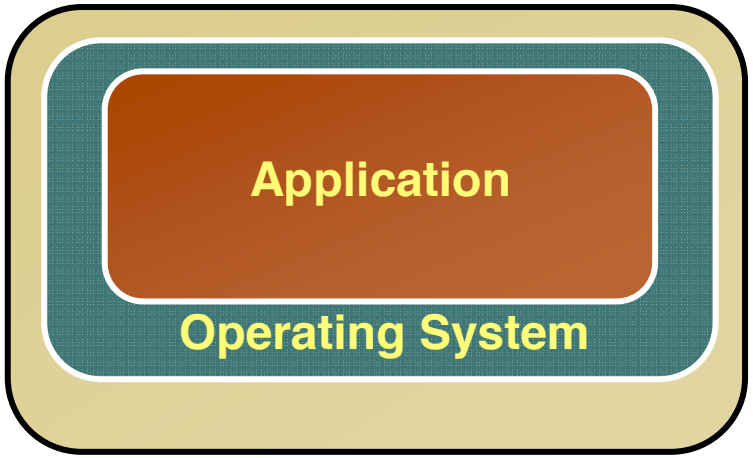
vmware®

Agenda

- Virtualization Today
- Technology Trends and the Future Datacenter
- Future directions
  - CPU Virtualization
  - I/O Virtualization
  - Virtual appliances
- Conclusions

vmware®

X86 Server Virtualization Basics

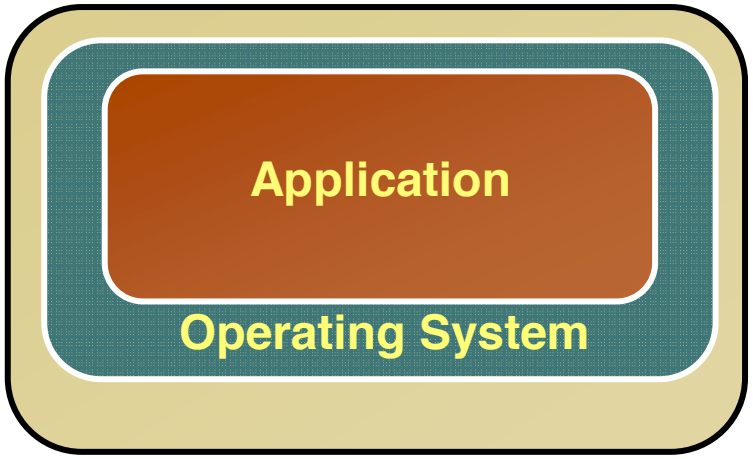


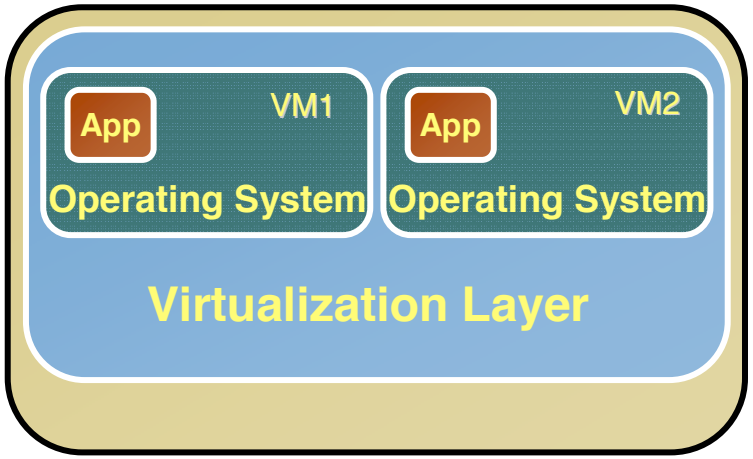
Before Server Virtualization:

- Single OS image per machine
- Software and hardware tightly coupled
- Running multiple applications on same machine often creates conflict
- Underutilized resources

vmware®

X86 Server Virtualization Basics





Before Server Virtualization:

- Single OS image per machine
- Software and hardware tightly coupled
- Running multiple applications on same machine often creates conflict
- Underutilized resources

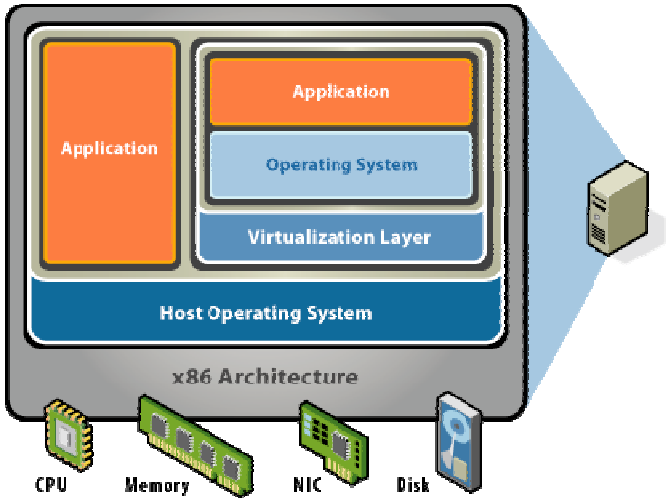
After Server Virtualization:

- Virtual machines (VMs) **break 1-to-1 dependency** between OS and HW
- Manage OS and application as single unit by **encapsulating** them into VMs
- Strong **isolation** between VMs
- **Hardware-independent:** they can be provisioned anywhere

vmware®

X86 Server Virtualization Architectures

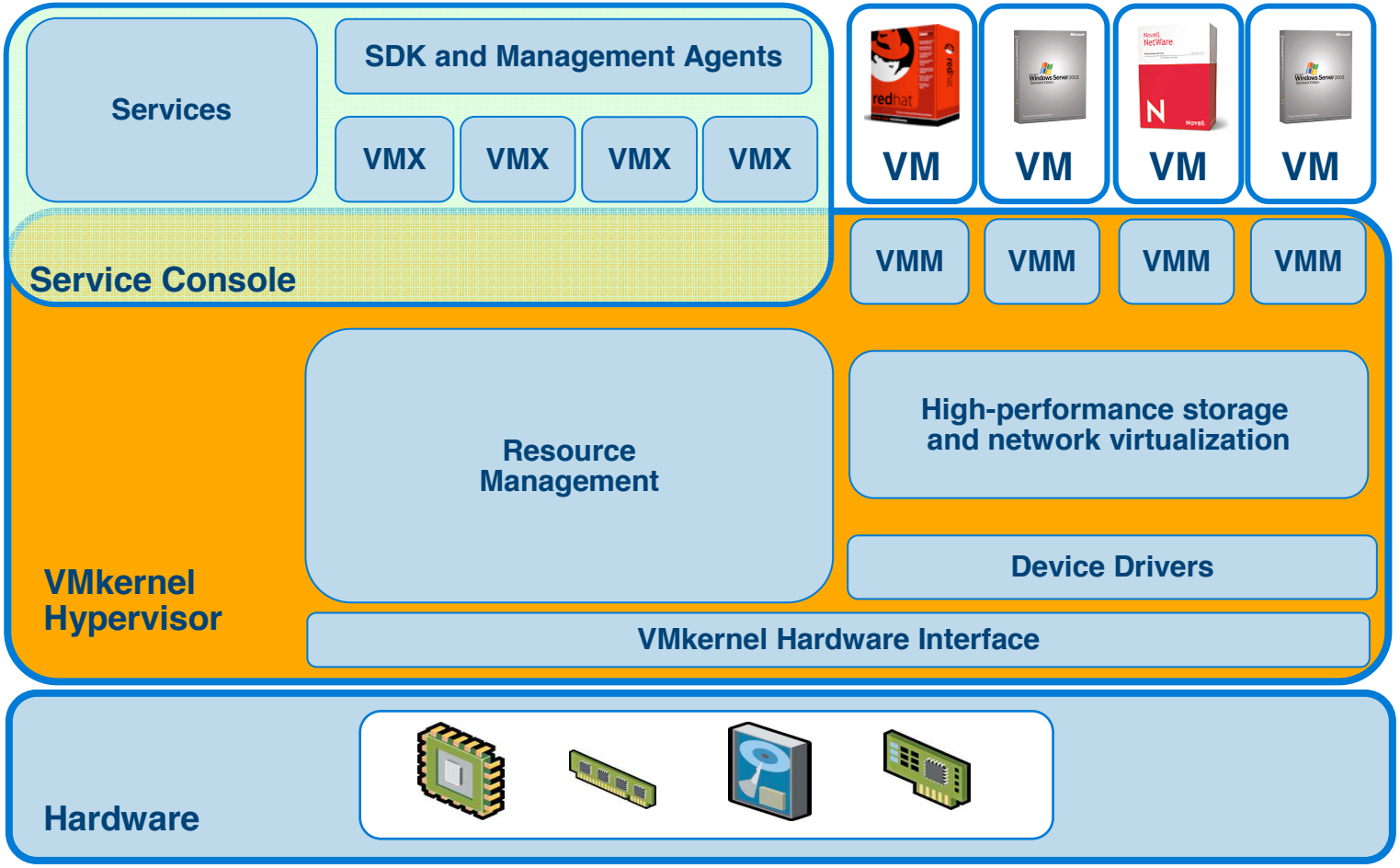
- Hosted Architectures
  - Install as applications on Windows/Linux with small context switching driver
  - Leverage host IO stack and resource management
  - Examples include VMware Workstation, VMware Server, Microsoft Virtual PC, Microsoft Virtual Server, …



- Bare-metal Architectures
  - "Hypervisor" installs directly on hardware
  - Approach acknowledged as direction for datacenter
  - VMware ESX Server, Xen, Microsoft Viridian

vmware®

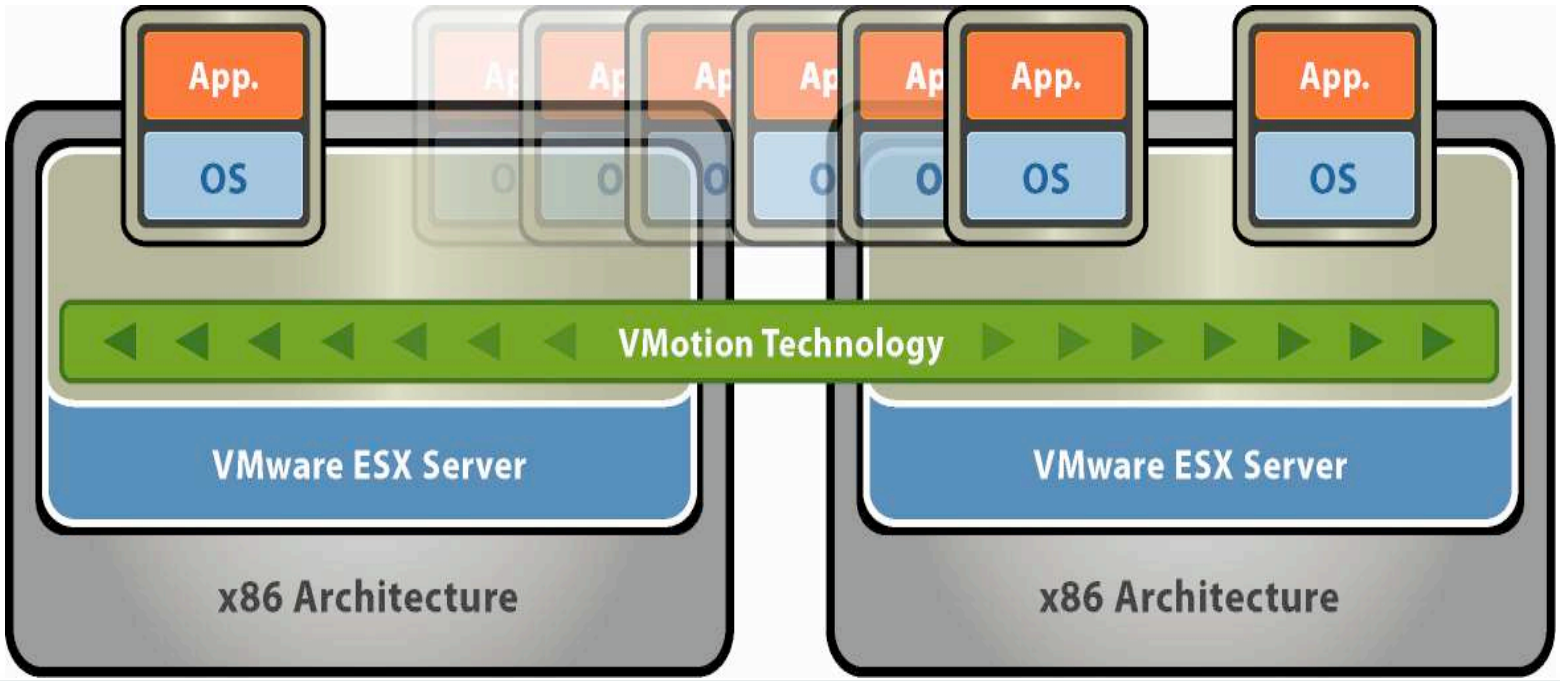
Bare-metal Example: VMware ESX Server



vmware®

Benefits Grow with Distributed Virtualization

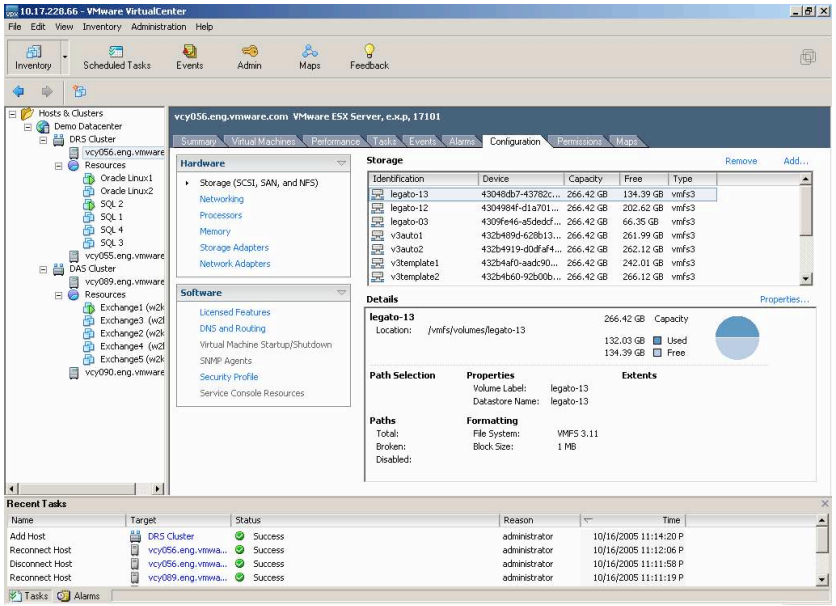
- Distributed file system allows multiple machines to see VMs
- Treat servers as a unified pool of resources
- Live migration (VMotion) of VMs between servers
  - Encapsulation and HW independence is key to this!

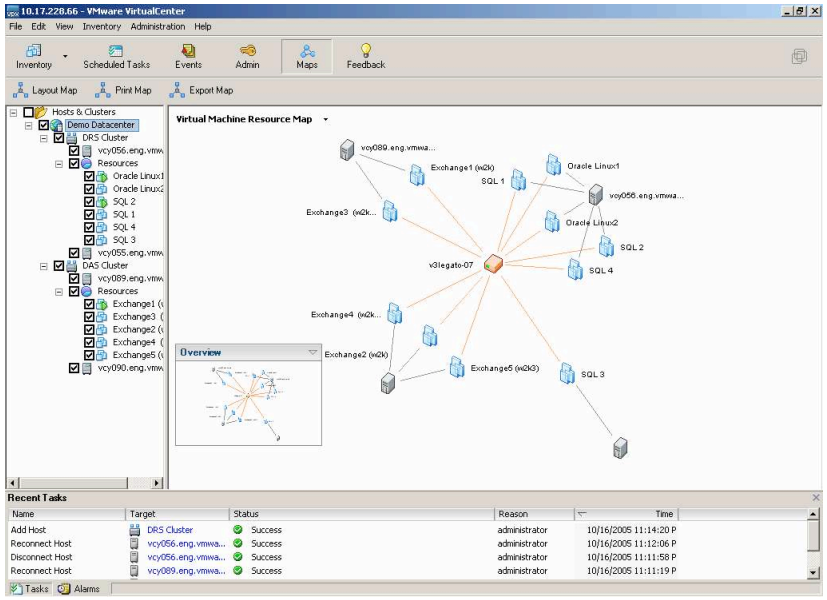


vmware®

Managing Distributed Virtualization

- Centralized management of hardware and VMs is key
  - Inventory of hardware and virtual machines (and their mappings)
  - Historic performance information
  - Remote consoles and devices
  - Drive VMotion between servers via drag-and-drop



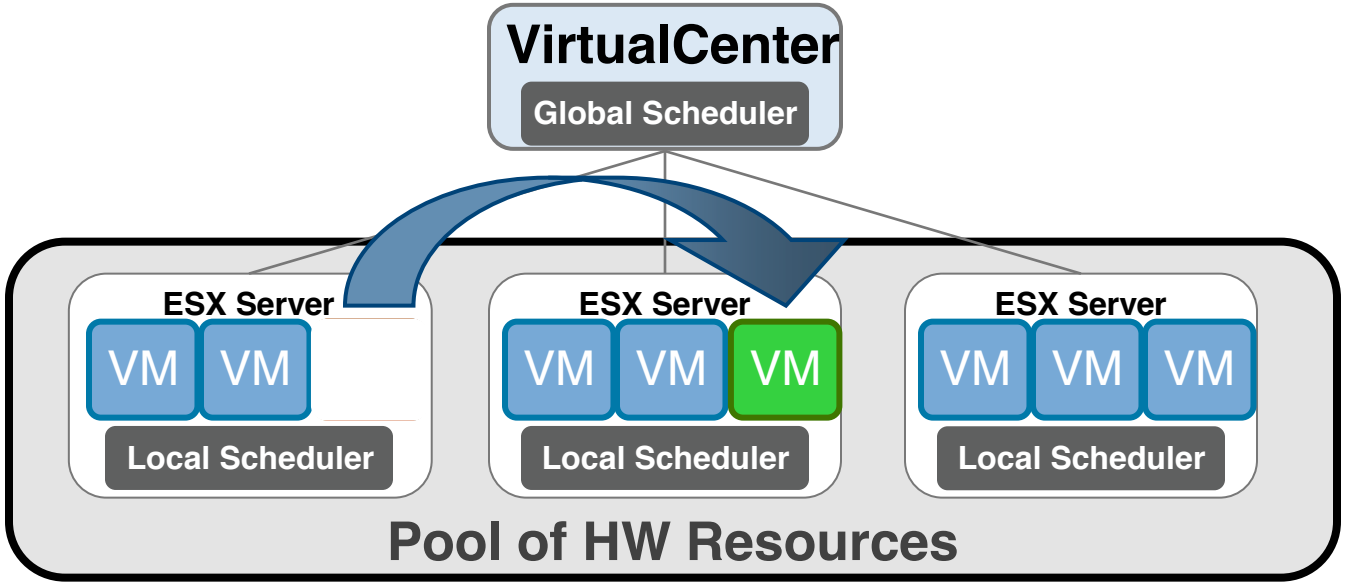


vmware

Distributed Virtualization Benefit: VMware DRS

With management, performance information, and VMotion:

- Input service level “rules” for each virtual machine
- Virtual Center uses VMotion to continuously optimize based on workload
- Reacts to adding or removing hosts from the cluster

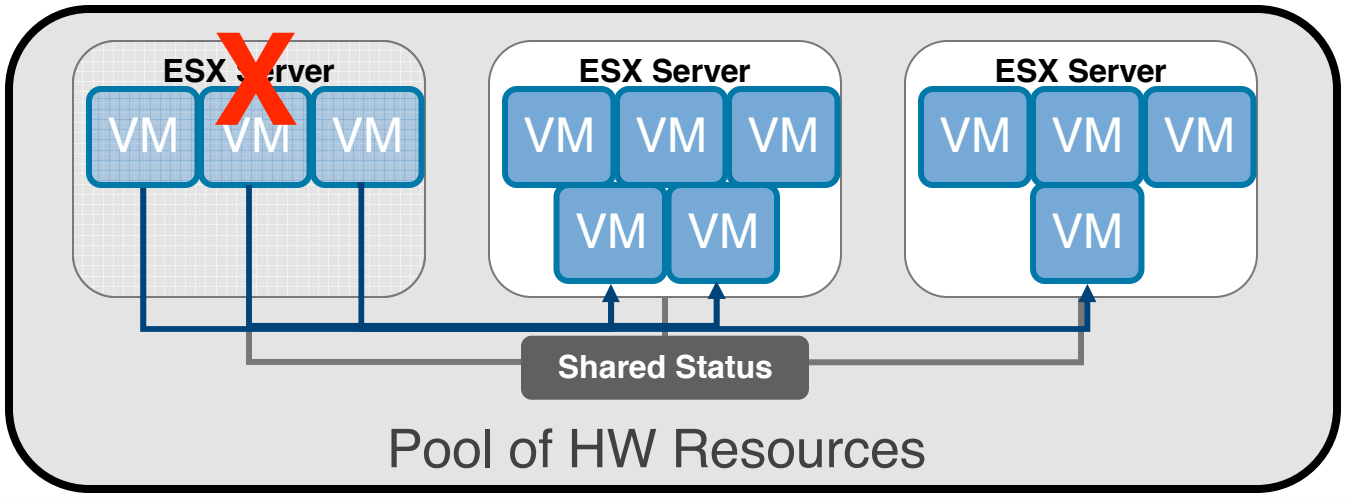


vmware®

Distributed Virtualization Benefit: VMware HA

With management, heartbeat, shared storage

- Losing a physical server means fewer resources, not lost virtual machines
  - Impacted virtual machines are restarted on remaining hosts
  - Placement optimized by global scheduler



vmware®

Common Virtualization Uses Today

Server Consolidation and Containment – Eliminate server sprawl by deploying systems into virtual machines

Infrastructure Provisioning – Reduce the time for provisioning new infrastructure to minutes with sophisticated automation capabilities. Like copying a file!

Business Continuity – Reduce the cost and complexity of business continuity by encapsulating entire systems files that can be replicated and restored onto any target server

Test and Development – Rapidly provision and re-provision test and development servers; store libraries of pre-configured test machines

Enterprise Desktop – Secure unmanaged PCs. Alternatively, provide standardized enterprise desktop environments hosted on servers.

Legacy Application Re-hosting – Migrate legacy operating systems and software applications to virtual machines running on new hardware for better reliability

Agenda

- Virtualization Today
- Technology Trends and the Future Datacenter
- Future directions
  - CPU Virtualization
  - I/O Virtualization
  - Virtual appliances
- Conclusions

Technology Trends in the Datacenter

- Multi-core CPUs (ISCA Session #4)
  - 16+ CPUs/cores per server
  - Increasing NUMA-ness
- 64-bit addressing
  - Enables huge amounts of physical memory
- Cooling and power costs soaring (ISCA Session #3)
  - Power-aware CPUs, servers, and racks
- Converged I/O fabrics
  - Shared high-speed interface to network and storage
- Network-based, virtualized storage
  - Stateless servers with flexible I/O connections

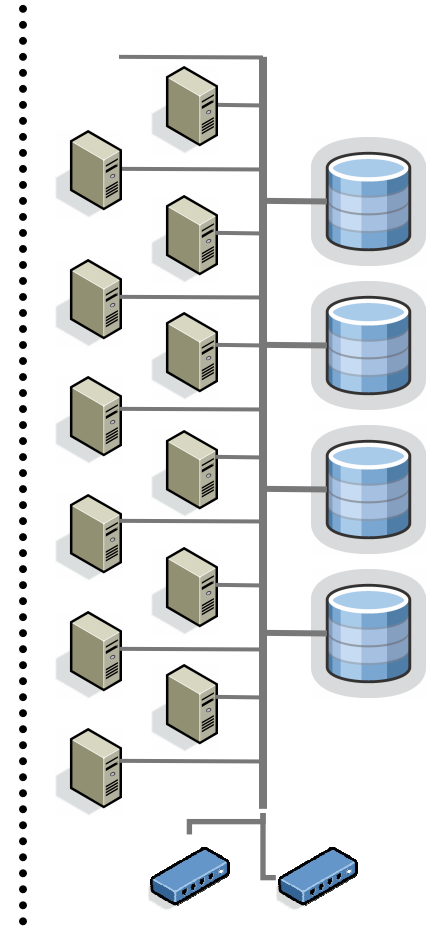
vmware®

Virtualization is Key to Exploiting Trends

- Allows most efficient use of the compute resources
  - Few apps take advantage of 16+ CPUs and huge memory as well as virtualization
  - Virtualization layer worries about NUMA, not apps
- Maximize performance per watt across all servers
  - Run VMs on minimal # of servers, shutting off the others
  - Automated, live migration critical:
    - Provide performance guarantees for dynamic workloads
    - Balance load to minimize number of active servers
- Stateless, Run-anywhere Capabilities
  - Shared network and storage allows flexible mappings
  - Enables additional availability guarantees

vmware®

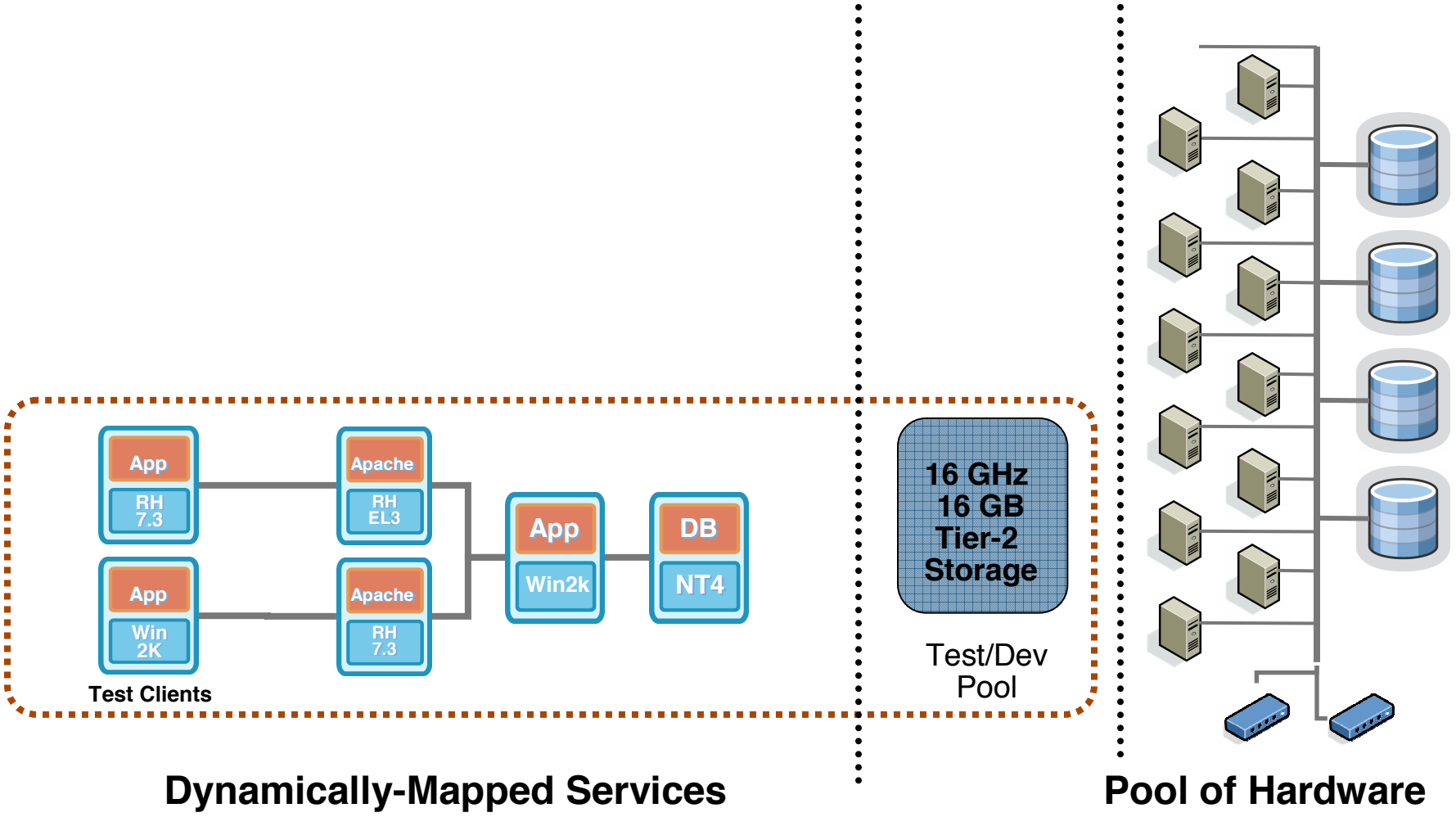
Vision: The Fully Virtual Datacenter



Pool of Hardware

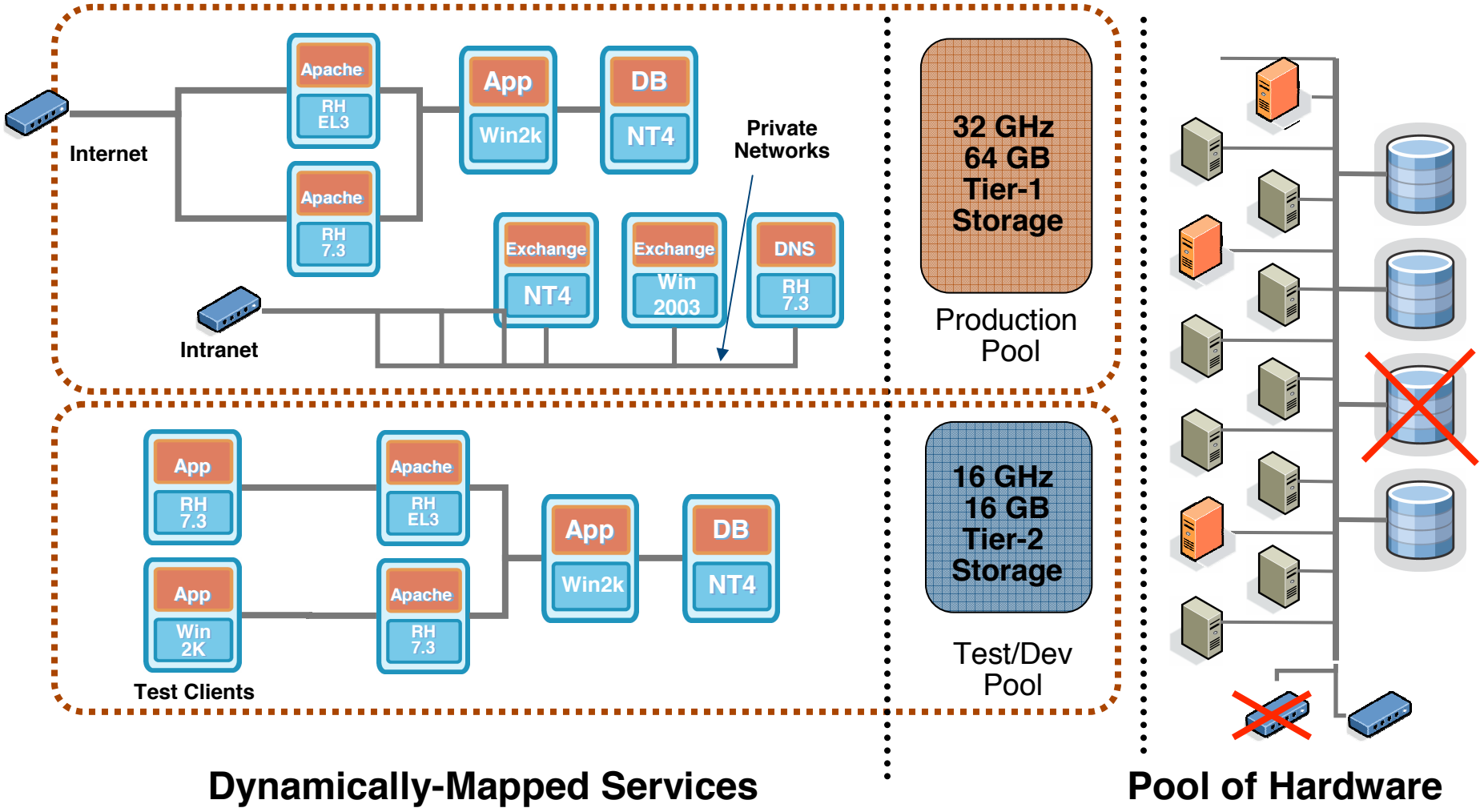
vmware®

Vision: The Fully Virtual Datacenter



vmware®

Vision: The Fully Virtual Datacenter



vmware®

Agenda

- Virtualization Today
- Technology Trends and the Future Datacenter
- Future directions
  - CPU Virtualization
  - I/O Virtualization
  - Virtual appliances
- Conclusions

vmware®

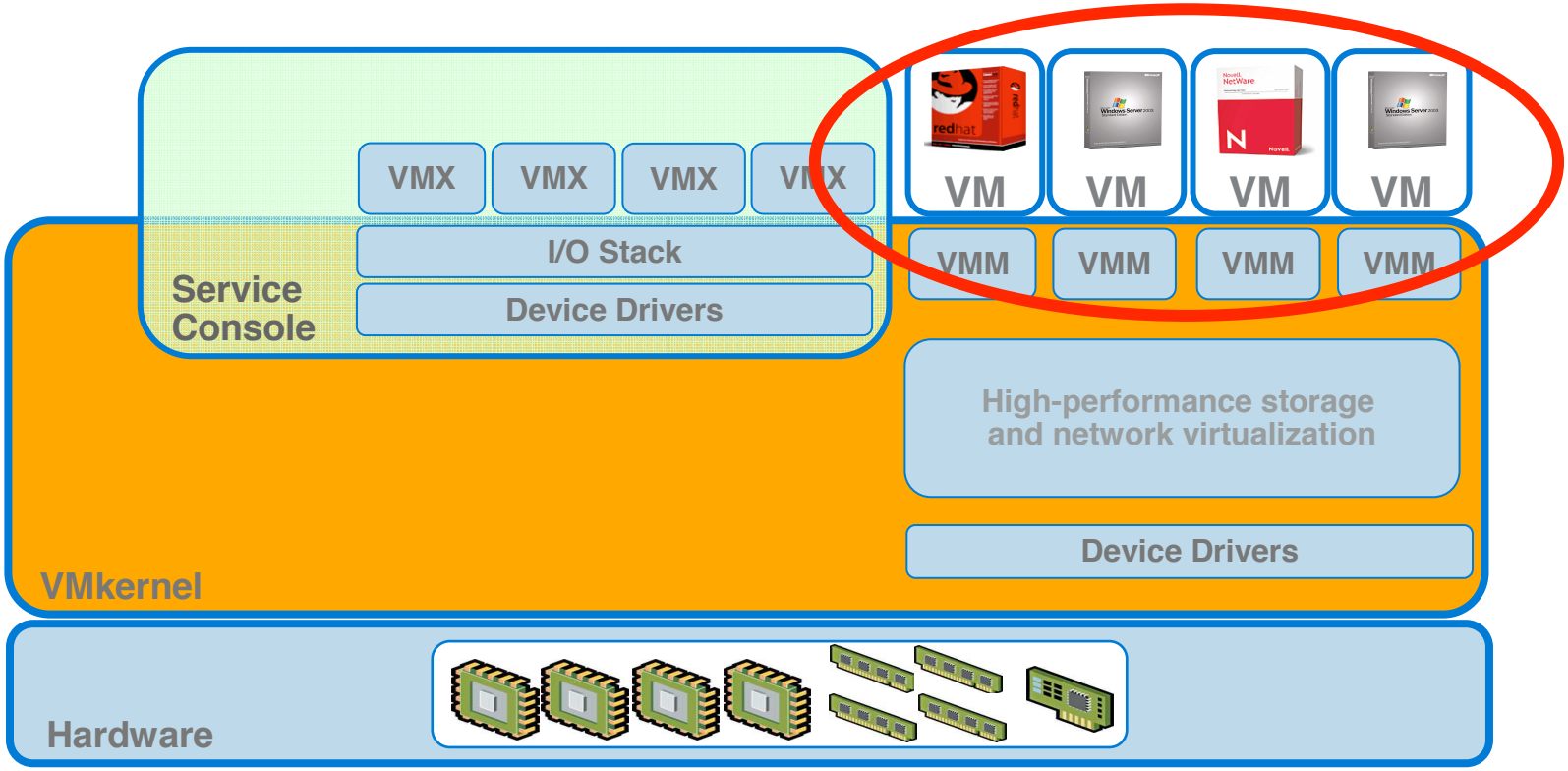
Hardware and OS Support on the Way

- New HW technologies provide explicit support
  - Increased CPU assist coming via Intel VT* and AMD SVM
    - Uber context switches
    - Better trapping
    - Hidden memory for hypervisor/vmm
  - I/O MMU's and virtualization-accelerating I/O devices
- Operating systems become virtualization-aware
  - Provide hints to the hypervisor about what is going on.
  - Skip expensive-to-virtualize operations altogether
  - Friendly licensing terms!

How do we best leverage this support?

vmware®

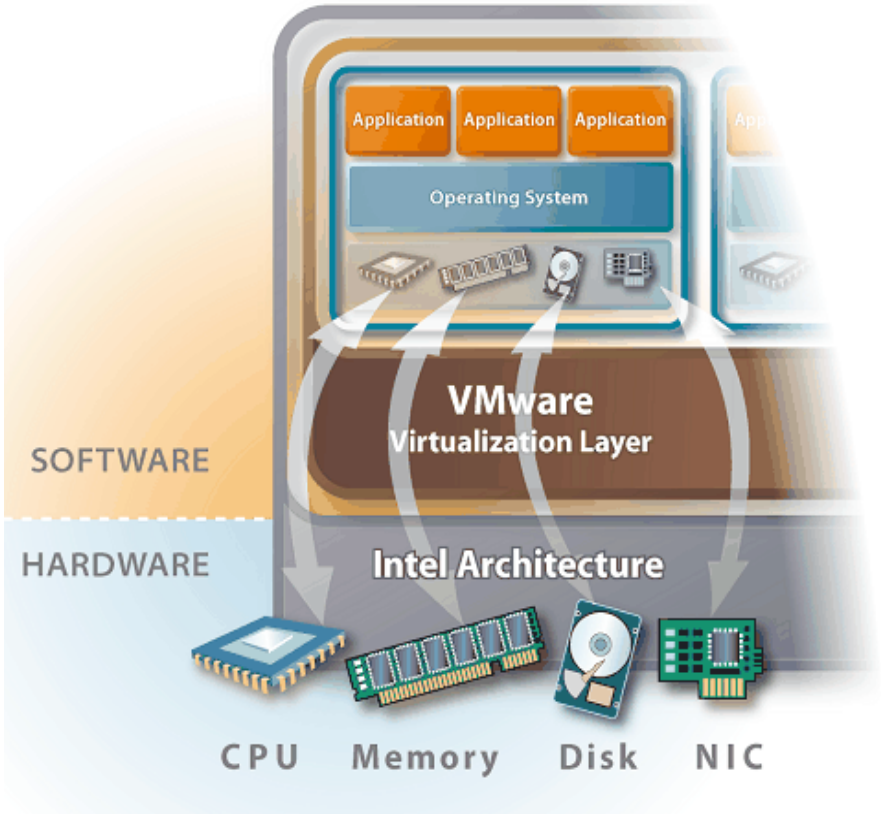
CPU Virtualization



vmware®

Virtual Machine Monitor (VMM)

- One for each VM
- Connects virtual resources to physical resources
- Must guarantee
  - Isolation of the VM
  - Compatibility
  - High performance



vmware®

Approach to CPU Virtualization

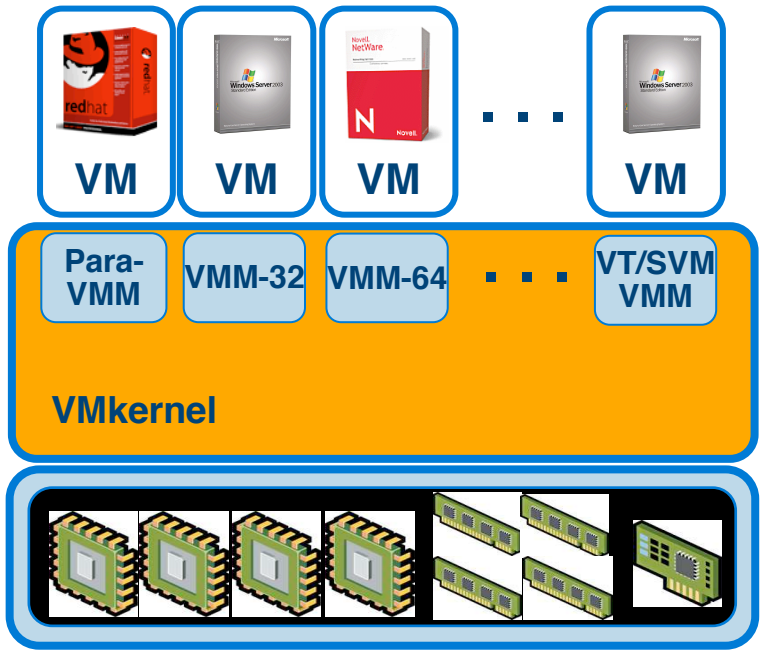
- Basic idea: directly execute code until not safe
- Handling unsafe code
  - Trap & emulate: classic approach, easier w/CPU support
  - Avoid unsafe code, call into VMM: paravirtualization
  - Dynamically transform to safe code: binary translation
- Tradeoffs among the methods

	Trap & Emulate	Para-virtualize	Binary Translate
Performance	**Average**	**Excellent**	**Good**
Compatibility	**Excellent**	**Poor**	**Excellent**
Implementation	**Average**	**Average**	**Hard**
Extra Capabilities	**Standard**	**Many**	**Many**

vmware®

Our Approach to CPU Virtualization

- We expect multiple, simultaneously-running VMMs
  - 32-bit BT VMM
  - 64-bit VT/SVM VMM
  - Paravirtualized VMM
- Use most efficient method for the hardware and guest type!



The best choices will change as HW support matures and paravirtualization APIs become standard

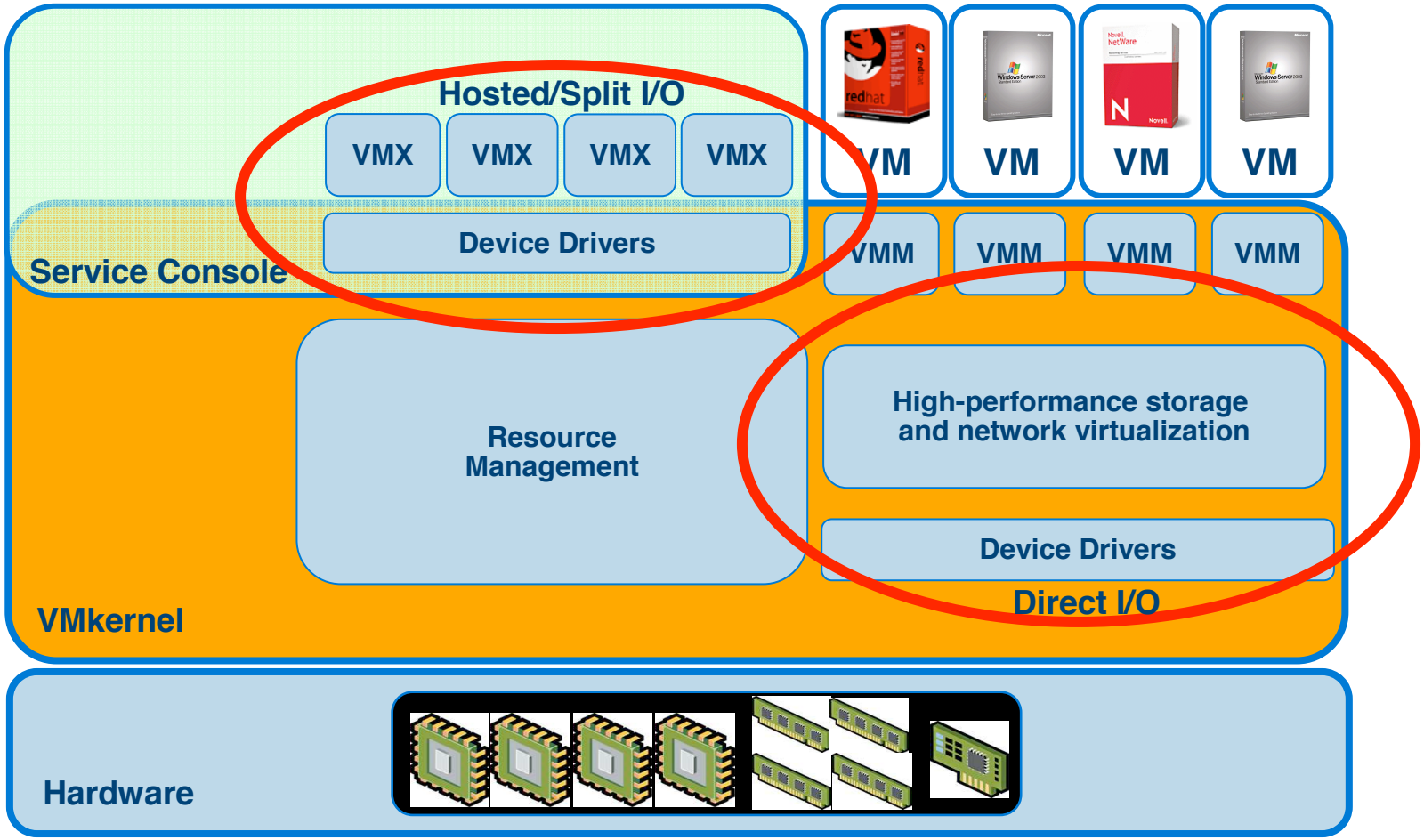
vmware®

Agenda

- Virtualization Today
- Technology Trends and the Future Datacenter
- Future directions
  - CPU Virtualization
  - I/O Virtualization
  - Virtual appliances
- Conclusions

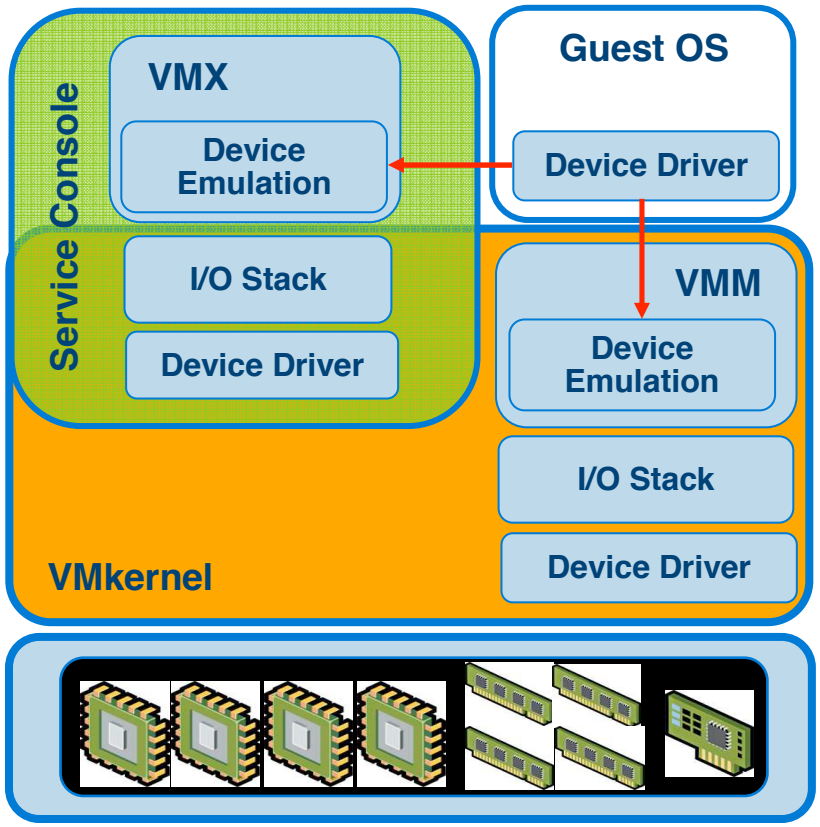
vmware®

I/O Virtualization



vmware®

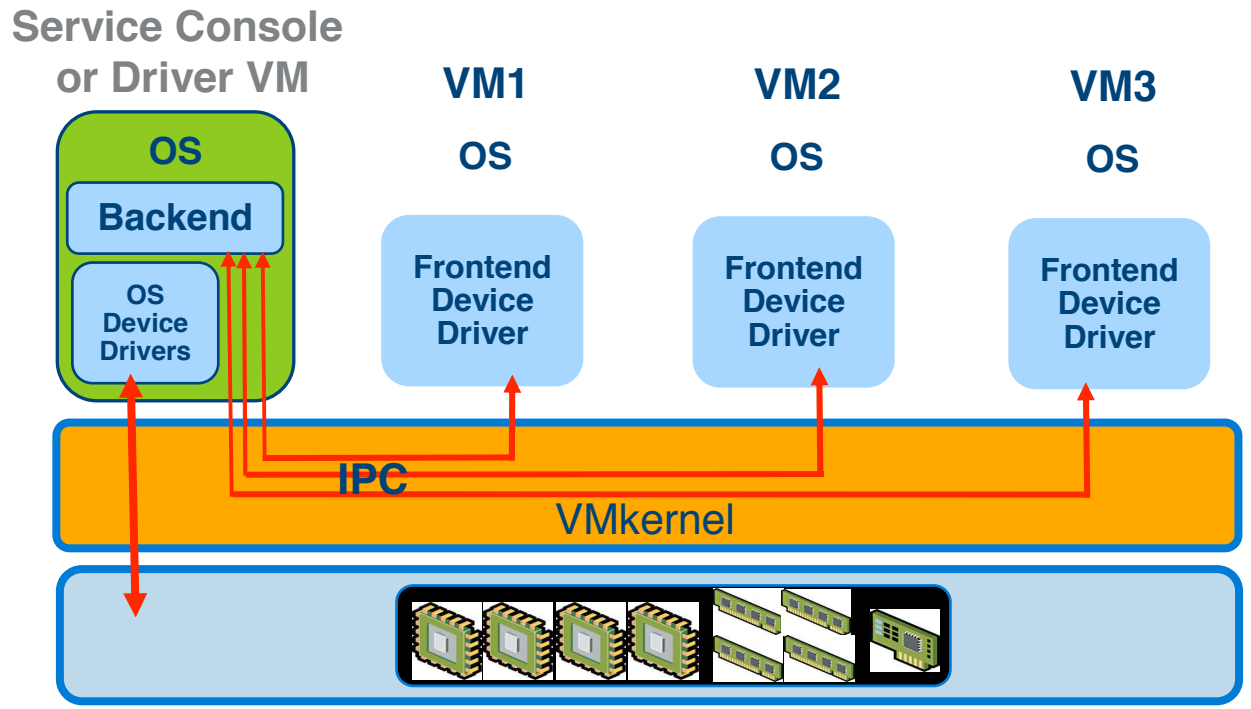
I/O Virtualization



- Full virtualization of I/O devices. Guests see standard devices regardless of physical machine
  - LSI SCSI Controller
  - E1000 NIC
- Virtual devices mapped to actual devices via
  - hosted/split I/O, or
  - direct I/O
- A third option in the future: passthrough

vmware®

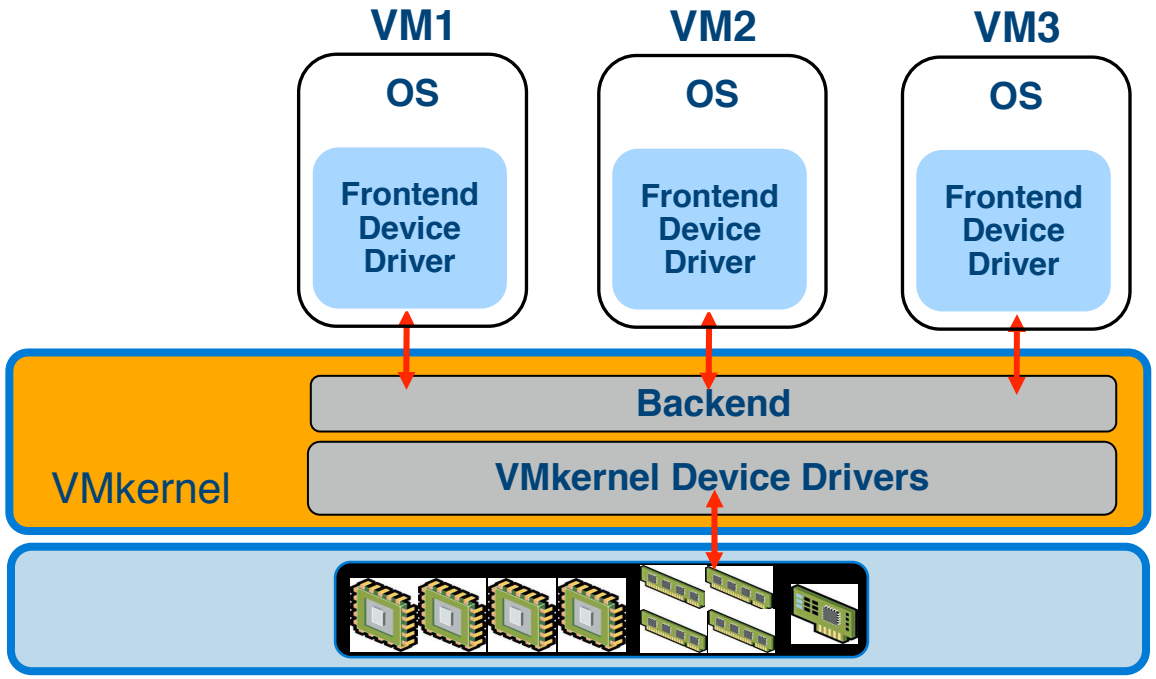
Hosted/Split I/O Virtualization



- IPC between frontend and backend
  - involves context switch and scheduling delays unless run on dedicated CPU
- Utilize drivers for Driver VM's OS type

vmware®

Direct I/O Virtualization



- System call between frontend and backend
  - Backend can run on any CPU
- Utilize VMkernel drivers
  - Linux compatibility layer (locks, memory allocation)

vmware®

Hosted/Split vs. Direct I/O

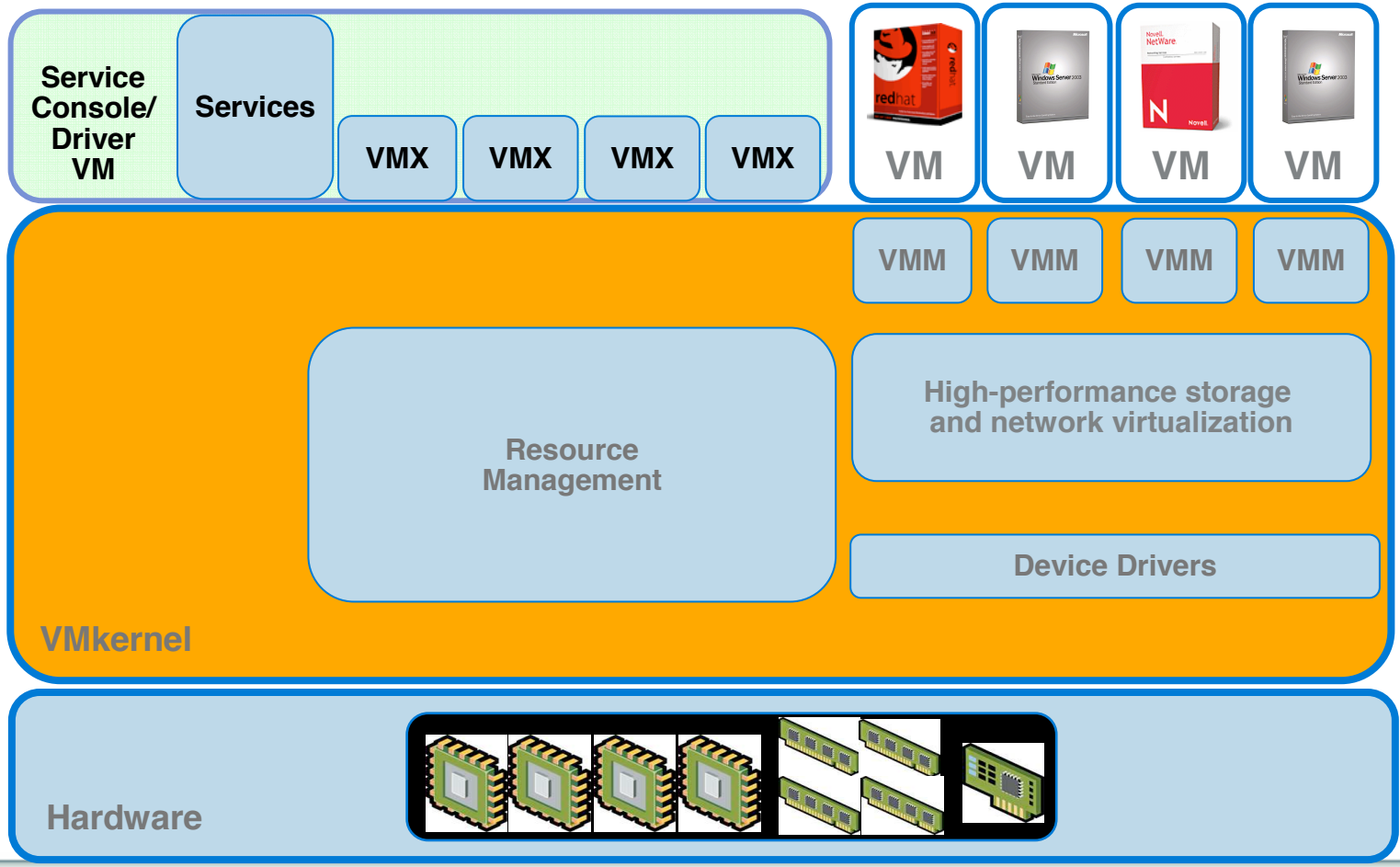
- Performance
  - Hosted/split uses IPC and generally requires dedicated CPUs
  - Direct I/O generally more efficient and scalable
- Compatibility
  - Hosted/split provides easier reuse of device drivers
  - Both require full qualification of drivers for unique workload patterns
- Fault Isolation
  - Hosted/split provides additional isolation with drivers in a VM
  - Direct I/O can take advantage of sandboxing and other techniques
  - Both require I/O MMU to provide true isolation

vmware®

I/O Models and Scalable performance

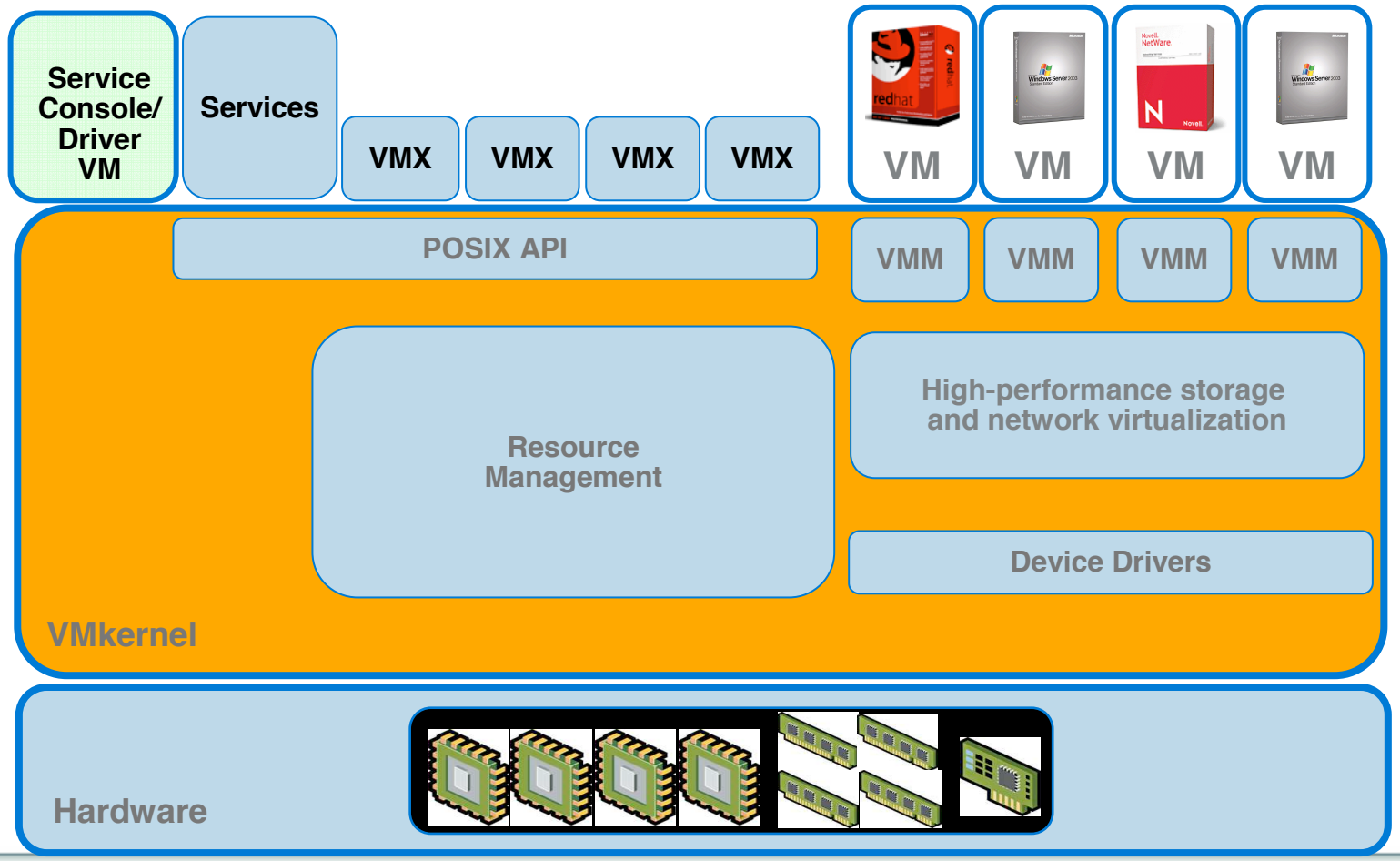
More Cores = More VMs

IO Backends in shared Service Console become bottleneck



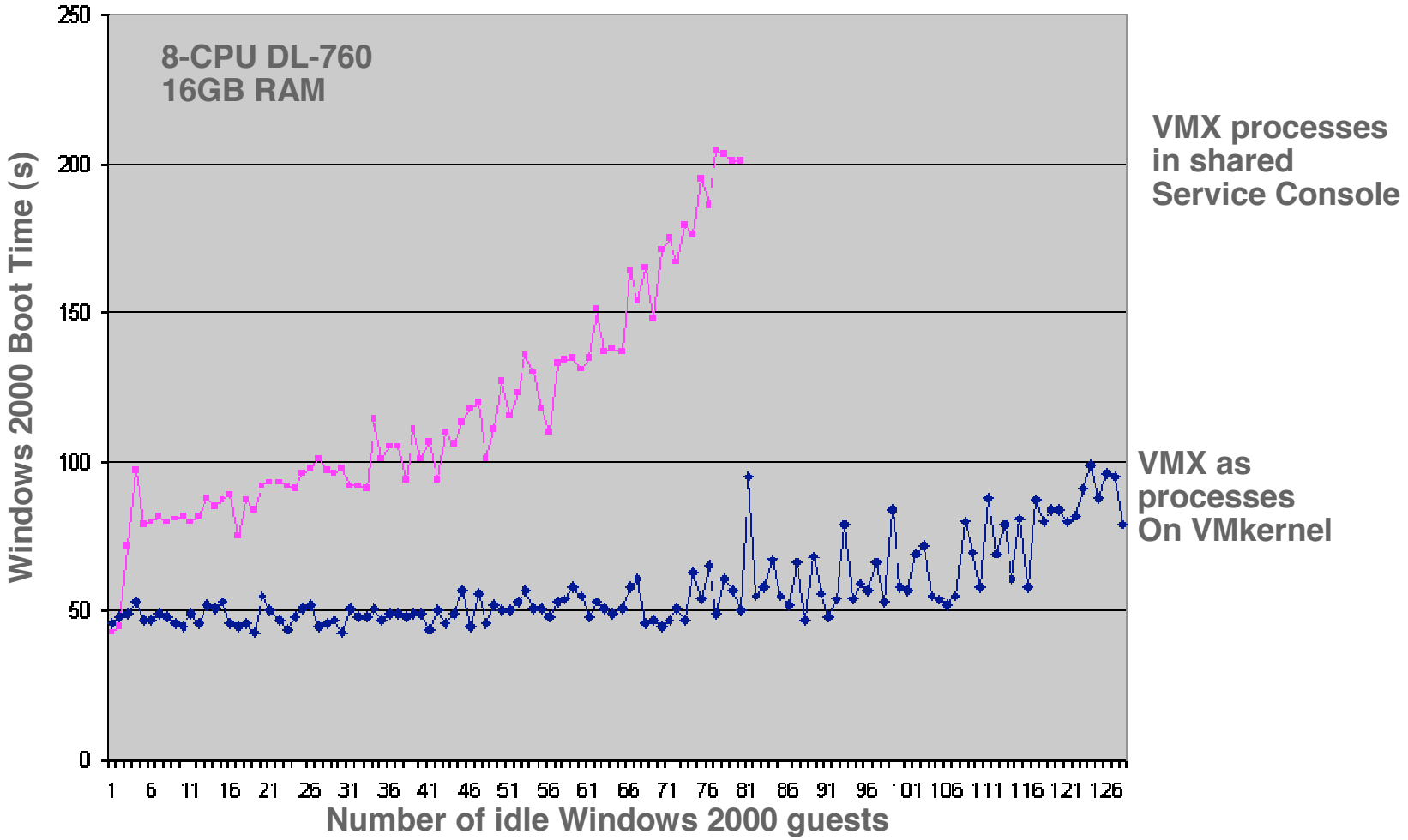
I/O Models and Scalable performance

Choice: Multiple service consoles or break out VMX



vmware®

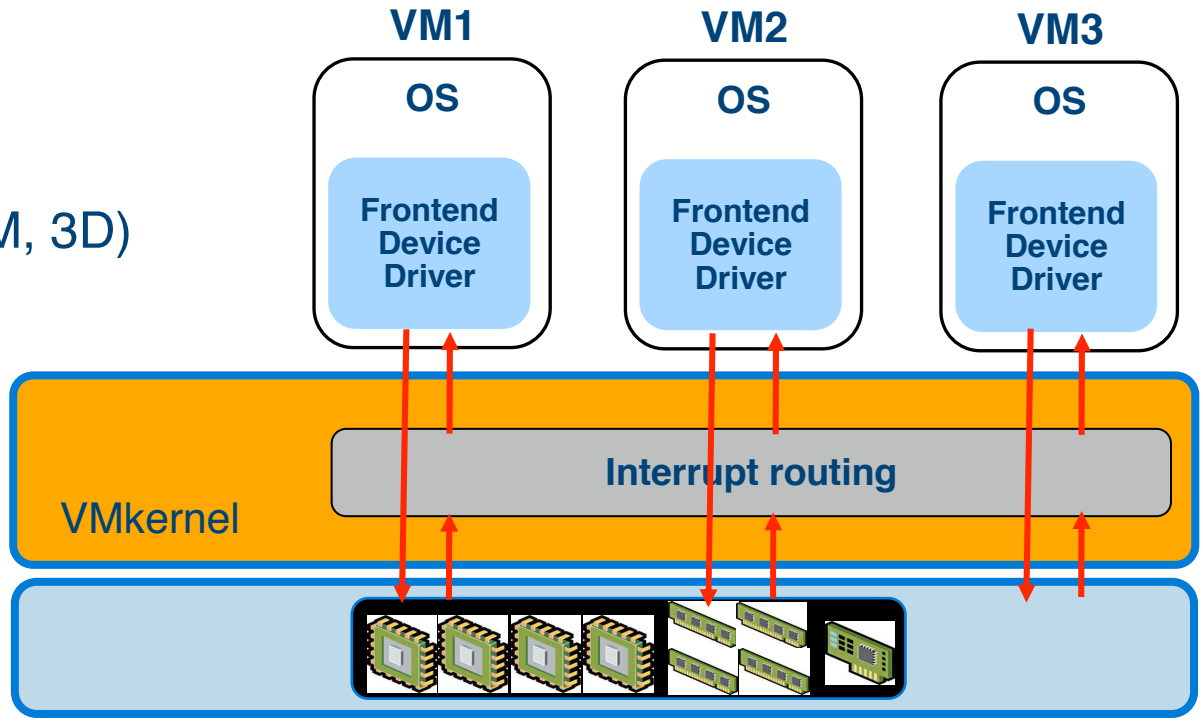
Scaling on 8-way Systems



vmware®

Passthrough I/O Virtualization

- Safely export hardware all the way to the guest
- Let guest OS driver directly drive the device
- VMkernel needed for set up and interrupt routing
- Some Use Cases
  - Performance (TOE)
  - Special HW (USB, TPM, 3D)
  - HW-level IO Isolation
- Challenges
  - Isolation
  - HW Independence



vmware®

Hardware Support for Passthrough

- To preserve capabilities, extra support required:
  - Isolation: I/O MMU to protect VMs from rogue DMA!
  - VMotion and Machine Independence:
    - Extracting and restoring device state (VMotion to same hardware)
    - Standardize device abstraction to VM (VMotion anywhere)
  - Memory over-commitment:
    - Device supports demand paging of memory it accesses
  - Device sharing:
    - Export multiple logical interfaces (e.g., WWNs via NPIV)
    - Track different I/O streams (e.g., tagged network queues)

vmware®

Summary of I/O Virtualization Tradeoffs

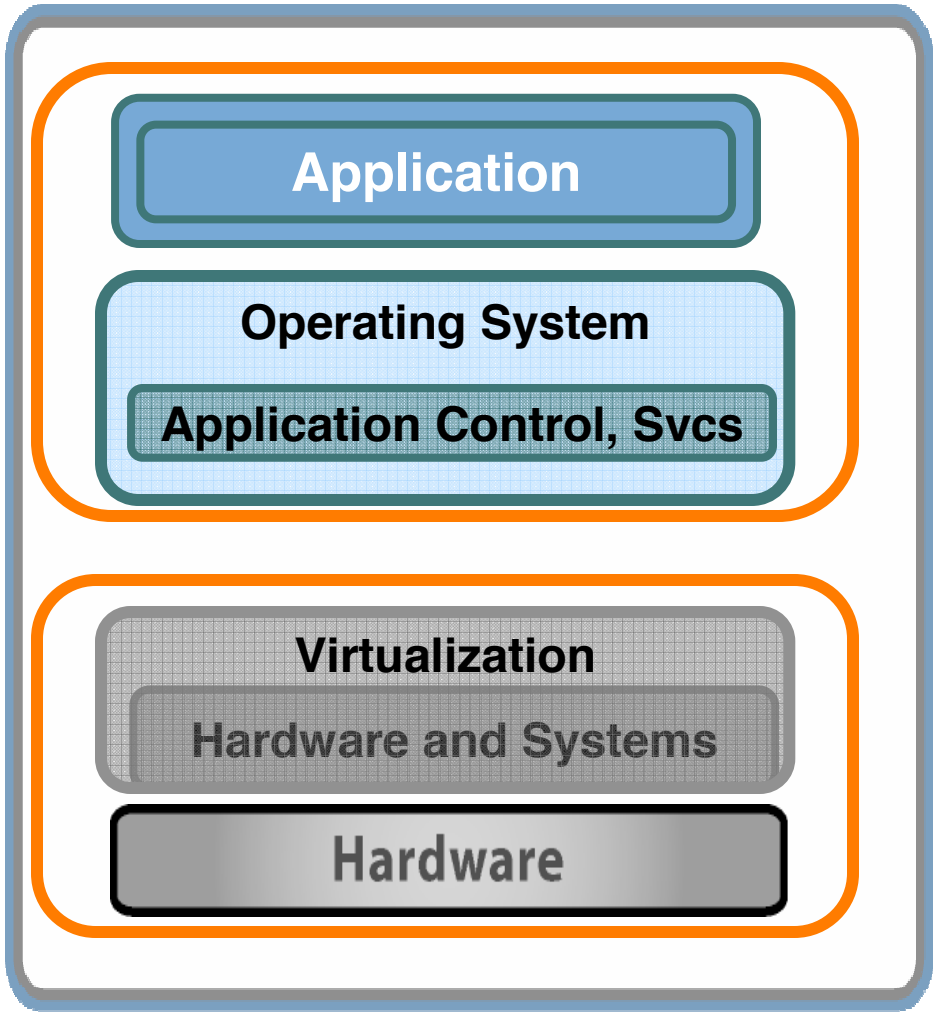
- Use hosted/split I/O for compatibility
  - Peak performance will require dedicated CPUs (Power)
  - Multiple Driver VMs for scalability
- Use direct I/O for top performance
  - Requires some driver porting
  - Improved fault isolation through use of I/O MMUs, sandboxing
- Use passthrough I/O for performance, fault isolation, specialized HW, and compatibility
  - Requires HW support to preserve virtualization functionality

vmware®

Agenda

- Virtualization Today
- Technology Trends and the Future Datacenter
- Future directions
  - CPU Virtualization
  - I/O Virtualization
  - Virtual appliances
- Conclusions

Virtualization Changes Traditional Boundaries



- OS Coupled With App
- Virtualization Coupled With Hardware

vmware®

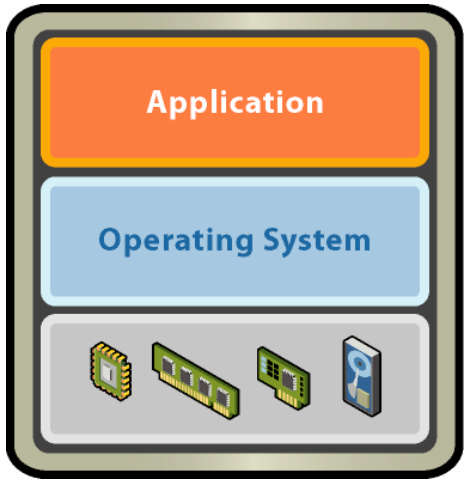
Question

If virtualization is everywhere, could these new boundaries inspire a new distribution model for software?

vmware®

Virtual Appliances

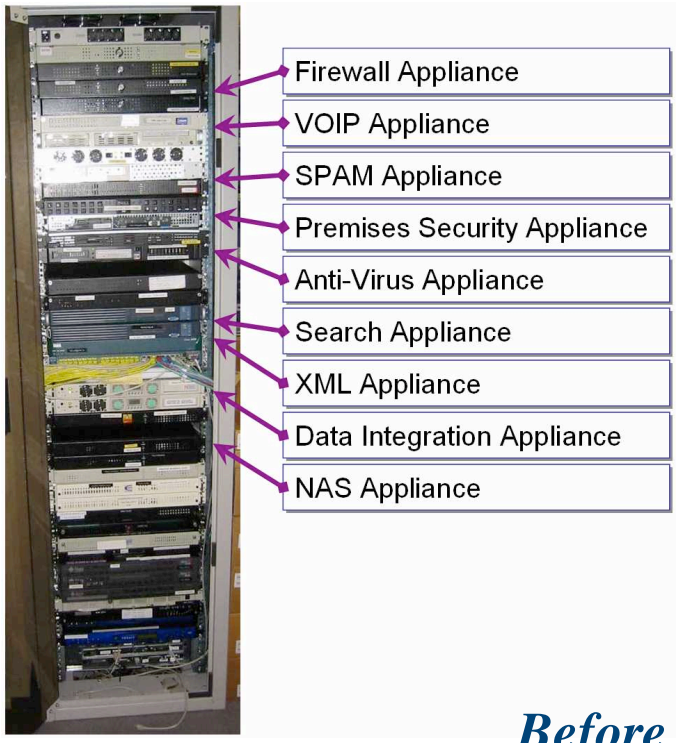
- **All the benefits of a traditional computing appliance without the cost and complexity**
- Pre-configured, purpose-built ***virtual*** device
- Pre-installed and pre-configured OS & application
- Limited configuration/customization exposed to user
- Simple installation and setup
- Doesn't require dedicated machine

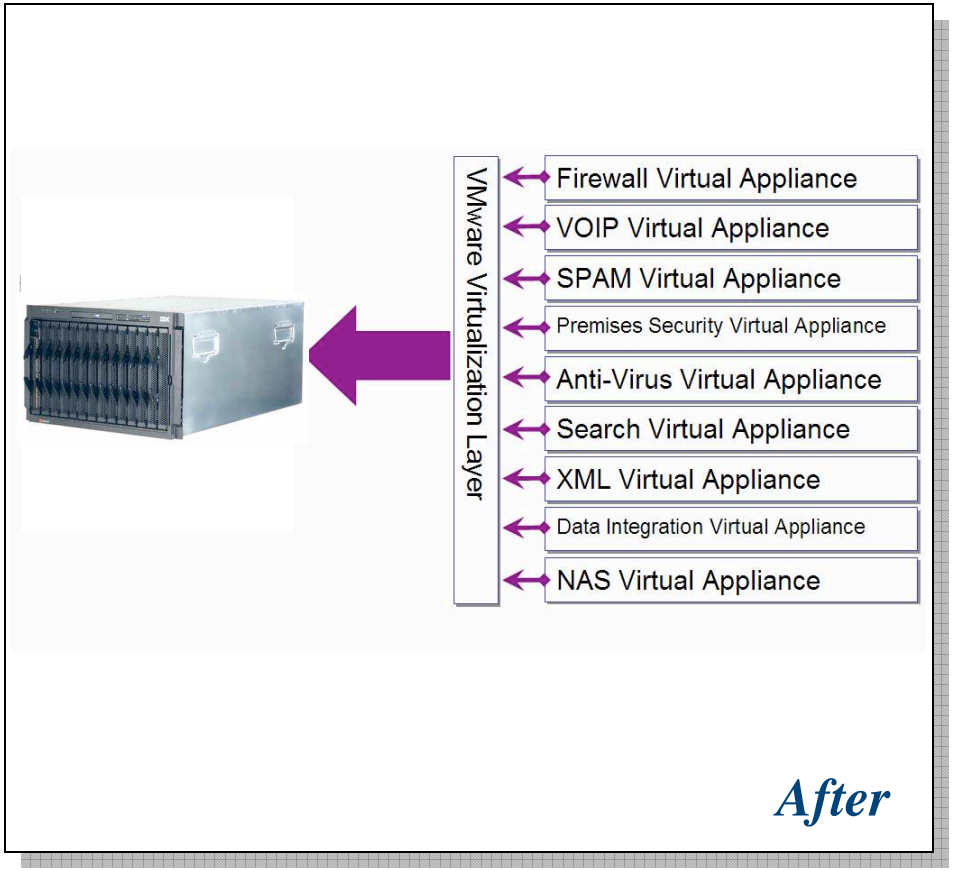


The Virtual Appliance approach provides a turn-key solution to complex software distribution

vmware®

Before / After Virtual Appliances





- One-to-one ratio of function to device
- Support from multiple sources
- Inefficient utilization of hardware

- Consolidate to save space/power
- Hardware support from preferred vendor
- Efficient utilization of hardware

vmware®

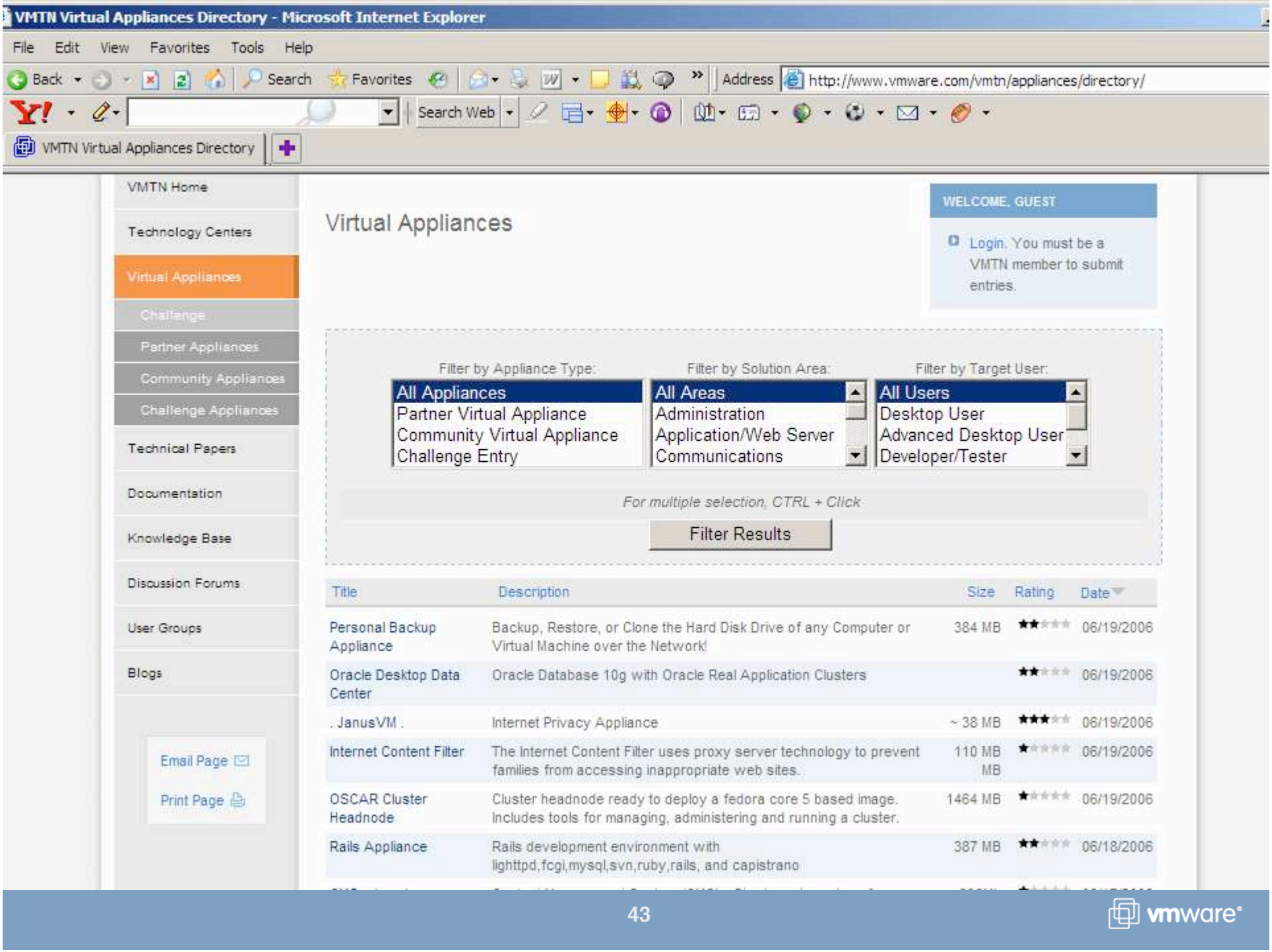
Benefits Matrix

	Traditional Software	Hardware Appliance	Virtual Appliance
Build and test on a controlled platform	✘ No	✓ Yes	✓ Yes
Simple plug-and-play installations	✘ No	✓ Yes	✓ Yes
Ability to tightly control access to underlying OS	✘ No	✓ Yes	✓ Yes
Inexpensive to distribute to customers	✓ Yes	✘ No	✓ Yes
Works with existing x86 hardware	✓ Yes	✘ No	✓ Yes
Low support cost and simple support logistics	✓ Yes	✘ No	✓ Yes
Easy, quick out-of-box experience for pilots, POC & Demonstrations	✘ No	✘ No	✓ Yes
Availability (easy backup and restore, easy recovery on new hardware, easy disaster recovery)	✘ No	✘ No	✓ Yes
Ability to scale up and down as needed without downtime	✘ No	✘ No	✓ Yes
Provides clustering capabilities without special code	✘ No	✘ No	✓ Yes

VM Library: Virtual Appliance Examples

- More than 200 submissions to recent contest!
- Browser Virtual Appliance
  - Very lean Ubuntu Linux installation + VM console access to Firefox
  - Internet browsing inside 'contained' environment of a Virtual Appliance
- Oracle RAC 10g on RHEL
  - Pre-installed and pre-configured to save time
  - Can see cluster behavior on a single machine
- Kid-safe computing
  - Web filter, replacement shell, squid URL blocking
- Voice Over IP (Asterisk@Home, sipX)

Free virtualization layer is a key enabler ala Adobe Reader


VMTN Virtual Appliances Directory - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Address http://www.vmware.com/vmtn/appliances/directory/
Search Web
VMTN Virtual Appliances Directory
VMTN Home
Technology Centers
Virtual Appliances
Challenge
Partner Appliances
Community Appliances
Challenge Appliances
Technical Papers
Documentation
Knowledge Base
Discussion Forums
User Groups
Blogs
Email Page
Print Page
Virtual Appliances
WELCOME, GUEST
Login. You must be a VMTN member to submit entries.
Filter by Appliance Type:
All Appliances
Partner Virtual Appliance
Community Virtual Appliance
Challenge Entry
Filter by Solution Area:
All Areas
Administration
Application/Web Server
Communications
Filter by Target User:
All Users
Desktop User
Advanced Desktop User
Developer/Tester
For multiple selection, CTRL + Click
Filter Results
Title Description Size Rating Date
Personal Backup Appliance Backup, Restore, or Clone the Hard Disk Drive of any Computer or Virtual Machine over the Network! 384 MB 06/19/2006
Oracle Desktop Data Center Oracle Database 10g with Oracle Real Application Clusters 06/19/2006
. JanusVM . Internet Privacy Appliance ~ 38 MB 06/19/2006
Internet Content Filter The Internet Content Filter uses proxy server technology to prevent families from accessing inappropriate web sites. 110 MB MB 06/19/2006
OSCAR Cluster Headnode Cluster headnode ready to deploy a fedora core 5 based image. Includes tools for managing, administering and running a cluster. 1464 MB 06/19/2006
Rails Appliance Rails development environment with lighttpd,fcgi,mysql,svn,ruby,rails, and capistrano 387 MB 06/18/2006

vmware®

Summary

- Virtualization is key to the future datacenter
  - Hardware trends require it for efficiency
- Examined two key areas
  - CPU virtualization
  - I/O virtualization
- Introduced the concept of virtual appliances
  - A new means of software distribution

vmware®

Calls to Action

- CPU Vendors
  - Continue the progress! (nested page table support, reduced latencies)
- I/O Device Vendors
  - Build required hardware support for pass-through
  - Wide open space for research and start-ups
- Software Developers
  - Support open, standardized paravirtualization interfaces to avoid fragmentation and encourage hypervisor competition
  - Consider the benefits of virtual appliances

Everyone: Help deliver the benefits of virtualization to the world!

Thank You!

vmware®