IBM TRAINING

G51

Virtualization Comparison: System z, System p and VMware

Reed A. Mullen

IBM
SYSTEM z9 AND zSERIES EXPO
October 9 - 13, 2006

Orlando, FL

RETURN TO INDEX

IBM System z9 and zSeries Expo – Orlando, Florida

Virtualization Comparison: System z, System p, and VMware

October 2006

Reed A. Mullen
mullenra@us.ibm.com
IBM Systems and Technology Group

IBM Systems

Acknowledgements

- **Thanks to my colleagues for their contributions**
  - Rory Canellis, IBM
  - Jim Rymarczyk, IBM
  - Romney White, IBM

Agenda

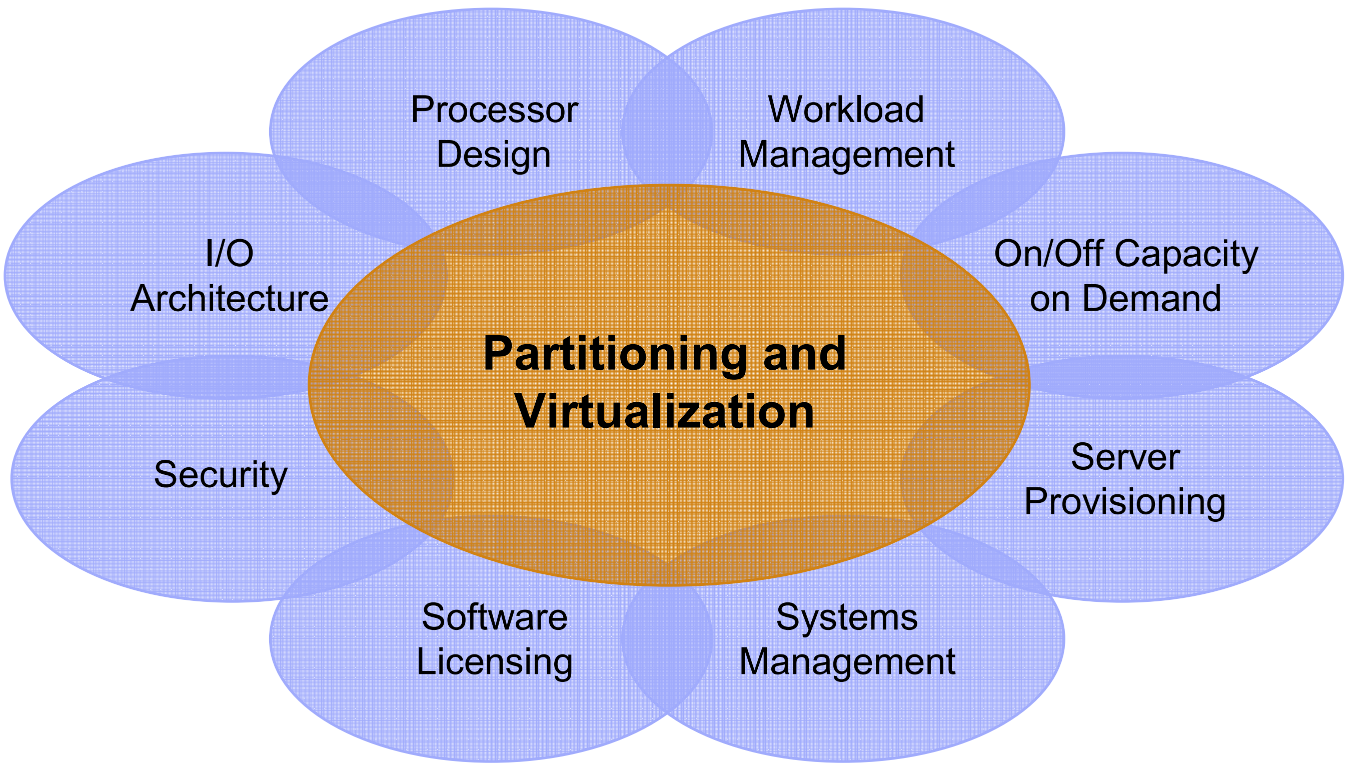
- **Server virtualization basics and business value**
- **Platform overview of virtualization technology**
  - Intel (VMware and others)
  - IBM System p (pSeries)
  - IBM System z (zSeries)
- **IBM System z virtualization differentiation**

Key Learning Points

- **Server virtualization support provides significant business value**
  - Reduces hardware expenses
  - Simplifies operational tasks
  - Facilitates on demand computing
- **Virtualization support is pervasive across server platforms, but the technologies and value propositions are not equal**
  - As with other server technologies, the IBM mainframe enjoys advantages in the area of virtualization as well:
    - High availability
    - Security
    - Highly granular resource sharing
    - Large-scale server hosting
- **Virtualization solutions for Intel architectures are growing fast in number**
  - The competitive landscape is expected to change in the coming months / years
  - VMware has dominated this space in recent years
  - Newcomers include Microsoft Virtual Server 2005, Xen, and Virtual Iron

Server Partitioning and Virtualization

An Important Consideration for All Areas of Server Design and Deployment



Partitioning Defined

- **Partitioning is the division of a single server's resources* into multiple, independent, isolated systems capable of running their own operating system**
- **Three types of partitioning:**
  - *Hardware* – resources are allocated to partitions on a one-to-one basis with the underlying physical hardware (no sharing among partitions)
  - *Logical* – resources are managed by hardware firmware and allocated to partitions with a finer granularity than hardware partitioning (resource sharing among partitions)
  - *Software* – resources are managed by a software layer, aggregated into shared resource pools, and apportioned to users as *virtual* system resources, separating the presentation of the resources from the actual physical entities

* Resources include: processors, memory, I/O adapters and devices, networking interfaces, co-processors

Hypervisor Technologies

“Trapping and Mapping” method

- Guest OS runs in user mode
- Hypervisor runs in privileged mode
- Privileged instructions issued by guest operating system(s) are trapped by hypervisor
- IA-32 (Intel) complications:
  - Some instructions behave differently in privileged and user modes
    - For example, “POPF” treatment of the interrupt enable flag
  - User mode instructions that access privileged resources/state cannot be trapped; instruction must be changed to something that can be trapped
- Some guest kernel binary translation may be required
- Originally used by mainframes in 1960s and 1970s (VM/370)
- Used today by VMware

Hypervisor Technologies

Hypervisor Call method ("Paravirtualization")

- Guest OS runs in privileged mode
- Hypervisor runs in super-privileged mode
- Guest OS kernel (e.g., AIX, i5/OS, Linux) is modified to do hypervisor calls for I/O, memory management, yield rest of time slice, etc.
- Memory mapping architecture is used to isolate guests from each other and to protect the hypervisor
- Used by POWER5 today

Hypervisor Technologies

Direct Hardware Support method

- Guest OS runs in privileged mode
- Guest OS can be run unmodified, but can issue some hypervisor calls to improve performance or capability
  - I/O (z/VM)
  - Yield time slice (PR/SM™ and z/VM)
- Extensive hardware assists for hypervisor (virtual processor dispatching, I/O pass-through, memory partitioning, etc.)
- Used by System z and zSeries (PR/SM™ and z/VM)

Virtual Resources

- Proxies for real resources: same interfaces/functions, different attributes.
- May be part of a physical resource or multiple physical resources.

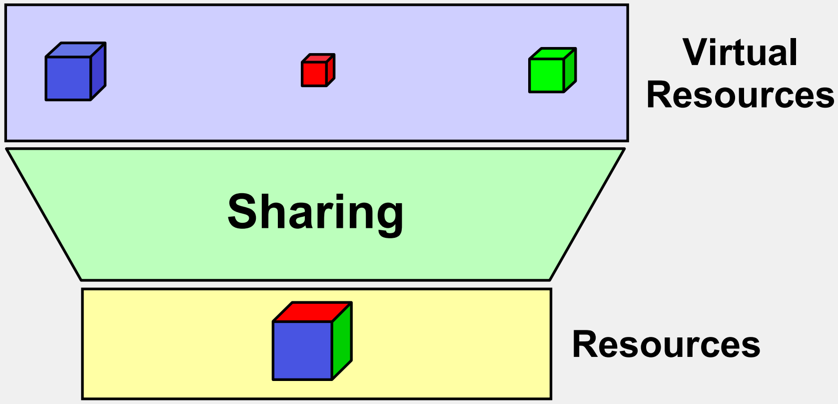
Virtualization

- Creates virtual resources and "maps" them to real resources.
- Primarily accomplished with software and/or firmware.

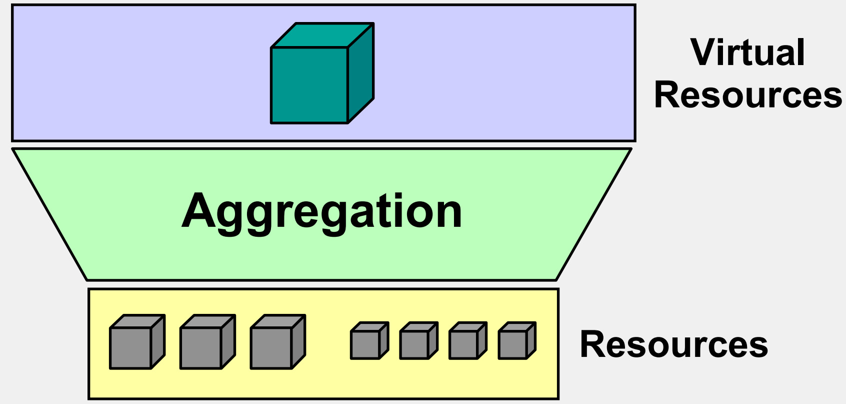
Resources

- Components with architected interfaces/functions.
- May be centralized or distributed. Usually physical.
- Examples: memory, disk drives, networks, servers.

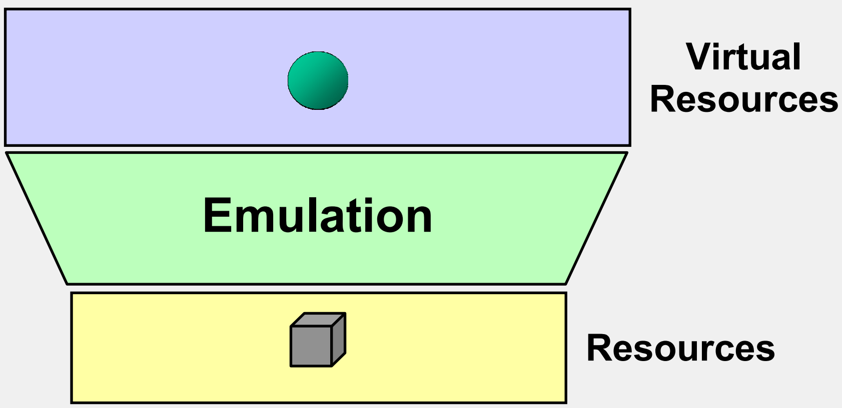
- **Separates presentation of resources to users from actual resources**
- **Aggregates pools of resources for allocation to users as virtual resources**



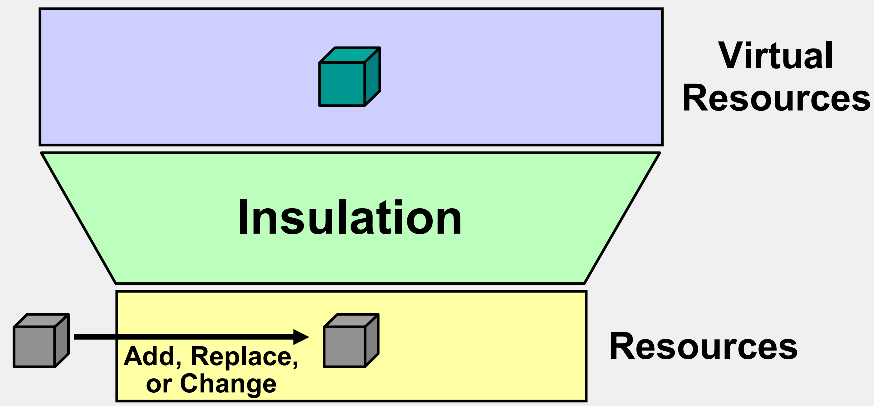
Examples: LPARs, VMs, virtual disks, VLANs
Benefits: Resource utilization, workload manageability, flexibility, isolation



Examples: Virtual disks, IP routing to clones
Benefits: Management simplification, investment protection, scalability



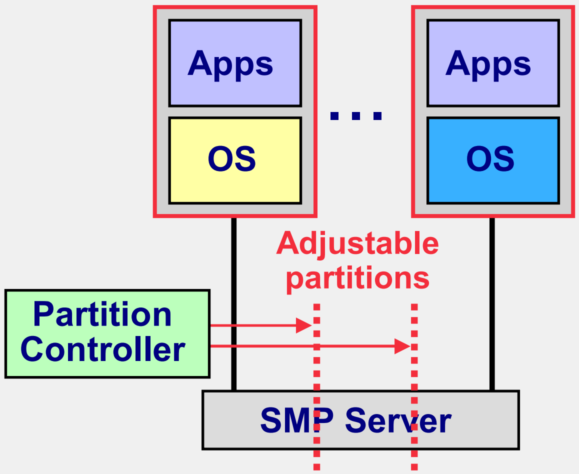
Examples: Arch. emulators, iSCSI, virtual tape
Benefits: Compatibility, investment protection, interoperability, flexibility



Examples: Spare CPU subst., CUoD, SAN-VC
Benefits: Continuous availability, flexibility, investment protection

Server Virtualization Approaches

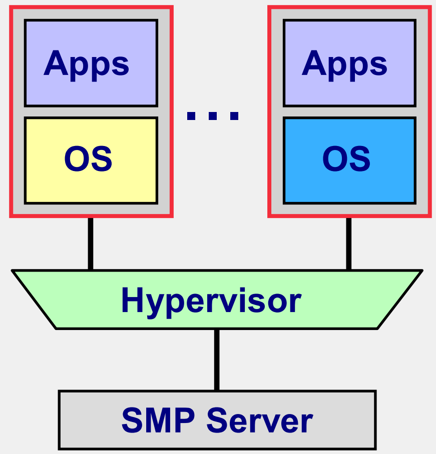
Hardware Partitioning



Physical partitioning
Sun Domains, HP nPartitions

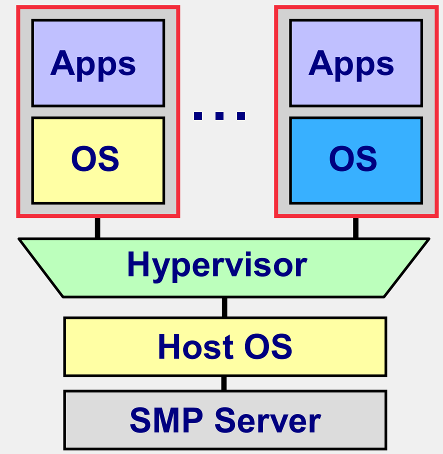
Logical partitioning
pSeries LPAR, HP vPartitions

Hypervisor: Type 1



Hypervisor software/firmware runs directly on server
zSeries PR/SM and z/VM
POWER Hypervisor
HP Integrity VM
VMware ESX Server
Xen Open Source Hypervisor
Virtual Iron, ScaleMP

Hypervisor: Type 2

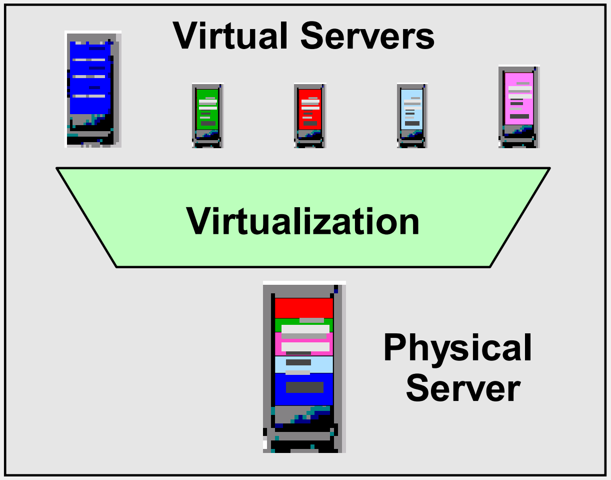


Hypervisor software runs on a host operating system
VMware GSX
Microsoft Virtual Server
Win4Lin
User Mode Linux

- Hardware partitioning subdivides a server into fractions, each of which can run an OS
- Hypervisors use a thin layer of code to achieve fine-grained, dynamic resource sharing
- Type 1 hypervisors with high efficiency and availability will become dominant for servers
- Type 2 hypervisors will be mainly for clients where host OS integration is desirable

Server Virtualization Business Value

Roles:
- Consolidations
- Dynamic provisioning/hosting
- Workload management
- Workload isolation
- Software release migration
- Mixed production and test
- Mixed OS types/releases
- Reconfigurable clusters
- Low-cost backup servers



Benefits:
- Higher resource utilization
- Greater usage flexibility
- Improved workload QoS
- Higher availability / security
- Lower cost of availability
- Lower management costs
- Improved interoperability
- Legacy compatibility
- Investment protection

In the final analysis, the virtualization benefits take three forms:

- **Reduced hardware costs**
  - Higher physical resource utilization
  - Smaller footprints
- **Improved flexibility and responsiveness**
  - Virtual resources can be adjusted dynamically to meet new or changing needs and to optimize service level achievement
  - Virtualization is a key enabler of on demand operating environments
- **Reduced management costs**
  - Fewer physical servers to manage
  - Many common management tasks become much easier

Platform Overview of Virtualization Technology

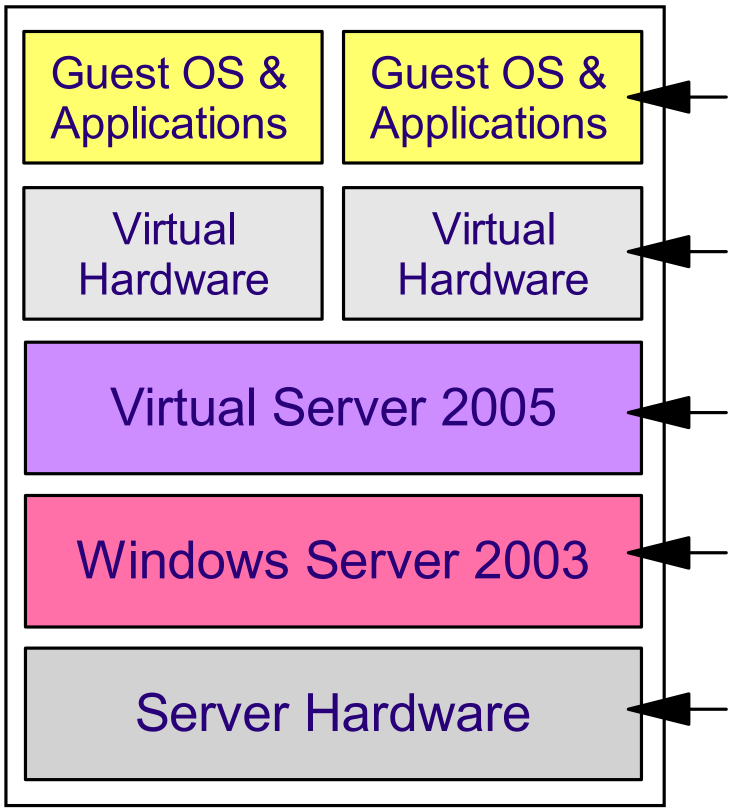
Surveying the Alternatives

- **Intel and AMD Software Solutions**
  - Microsoft Virtual Server 2005
  - VMware ESX Server
  - Xen
  - Virtual Iron
  - Others (e.g., SWsoft Virtuozzo, QEMU, …)
- **Hardware architecture enhancements are arriving now…**
  - Intel: Virtualization Technology (Silverdale/Vanderpool) – “With enhancements to Intel’s various platforms, Intel Virtualization Technology can improve the robustness and performance of today’s software-only solutions.”
  - AMD: Pacifica – “Pacifica is designed to provide foundation technologies to deliver IT resource utilization advantages through server consolidation, legacy migration and increased security.”

Microsoft Virtual Server 2005

- **Allows multiple operating systems to run simultaneously on the same processor**
- **Each independent virtual machine functions as a self-contained computer**
- **Virtual machines are encapsulated in portable Virtual Hard Disks (VHDs)**
  - Up to 32 VHDs can connect to a single virtual machine
  - VHDs can expand as data is added and "differenced"
  - VHD format is available royalty free
- **Virtual networking options are available**
- **Requires a hosting operating system (Windows Server 2003)**
- **Initially intended for hosting unsupported Windows NT environments**
- **Native Windows Server 2003 environment still recommended for many workload deployments**
- **Add-ins are available to support Linux**
  - Must be installed in the Linux guest images
  - Designed to improve guest-host interactions
- **Virtual Server 2005 is available as a free download**

Microsoft Virtual Server 2005 - Architecture



Windows NT 4.0 Server, Windows 2000 Server, Windows 2003 Server supported

1 CPU per virtual machine
Up to 3.6 GB of memory per virtual machine

Multithreaded virtual machine monitor (VMM) provides isolation

Windows Server 2003 32-bit supported host
Broad device compatibility

Optimized for 2 to 8 way servers
Scales up to 32 CPUs and up to 64 GB

Support for Intel and AMD virtualization technologies targeted for early 2007

Intel Partitioning with VMware ESX Server 3

Part of VMware Infrastructure 3

- **VMware ESX Server 3 runs directly on the hardware**
  - No hosting operating system required (as is the case with VMware GSX Server)
- **Creates multiple virtual machines on a single Intel system**
  - Supports a maximum of 128 virtual machines per VMware image (depending on system resources)
- **Manages resource allocations among virtual machine images**
  - Strong fault and security isolation (uses CPU hardware protection)
  - Virtual networking support available (MAC or IP addressing)
  - Direct I/O passthrough
- **Shared data cluster-ready**
- **VMware File System (VMFS) allows multiple virtual disks to be stored on a single LUN or partition**
- **Virtual machines are encapsulated**
  - Which enables relocation of virtual machine from one copy of VMware to another
- **Add-on products available**
  - VMware Virtual SMP allows virtual machines to be configured with up to 4 CPUs
  - VMware VirtualCenter

VMware VMotion

- **A key component of VMware VirtualCenter**
- **Allows movement of live applications across distributed systems without server interruption**
- **Perform zero-downtime, rolling hardware upgrades**
- **Migrate virtual machines to new hosts when resource availability becomes an issue**

VMware VirtualCenter 2

Part of VMware Infrastructure 3

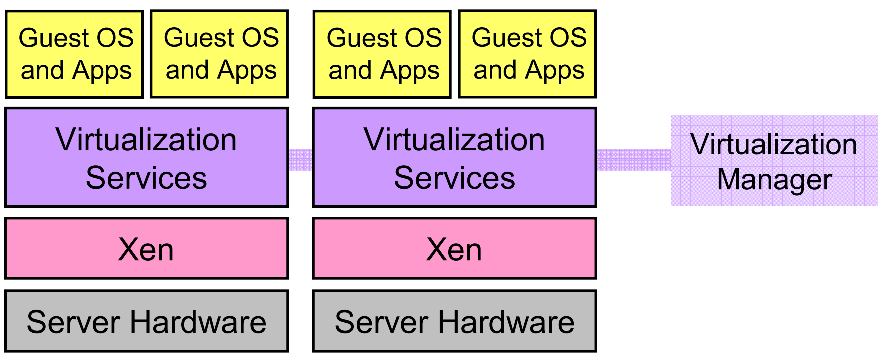
- **VirtualCenter Management Server**
  - The control node for configuring, provisioning and managing multiple VMware ESX servers
  - Runs on Windows 2000, Windows XP Pro, or Windows Server 2003
- **VirtualCenter Agent**
  - Needed to connect ESX servers with the Management Server
- **Virtual Infrastructure Client**
  - Allows users to connect to the VirtualCenter Management Server or individual ESX servers from any Windows PC
- **VirtualCenter Database**
  - Stores information used by the Management Server
  - Requires Oracle, Microsoft SQL Server, or Microsoft MSDE
- **Virtual Infrastructure Web Access**
  - Allows virtual machine management and access to graphical consoles without installing a client

Xen 3.0

- Open Source virtualization software solution based on Linux
- Uses paravirtualization to abstract CPU, memory, and I/O resources
- Guest operating systems are responsible for allocating and managing page tables
- Management and control software runs in Domain 0
- Intel Virtualization Technology (IVT) will enable hosting of unmodified guest operating systems
  - Windows support on Xen is not possible without IVT-capable hardware
- IBM is actively contributing to the Xen open source project

Virtual Iron 3.0

- **Requires Intel Virtualization Technology (IVT) and AMD Pacifica CPUs**
- **Replaces the proprietary hypervisor used in Virtual Iron 2.0 with Xen**
- **Virtual Iron Virtualization Services runs on top of Xen hypervisor**
  - Provides dynamic infrastructure management with hot-pluggable resources, server partitioning, and live migration of guest images
  - Supports Linux and Windows guest images (32- and 64-bit systems)
  - Up to 8 CPUs, 128GB RAM, 8 virtual NICs per virtual server
- **Virtual Iron Virtualization Manager provides command and control**
  - Java application that runs in any browser
  - Manage OS and application images (install, deploy, clone, etc.)
- **Virtual Iron 3.0 beta**
  - Available for Linux guests since July 2006
  - Available for Windows guests since September 2006



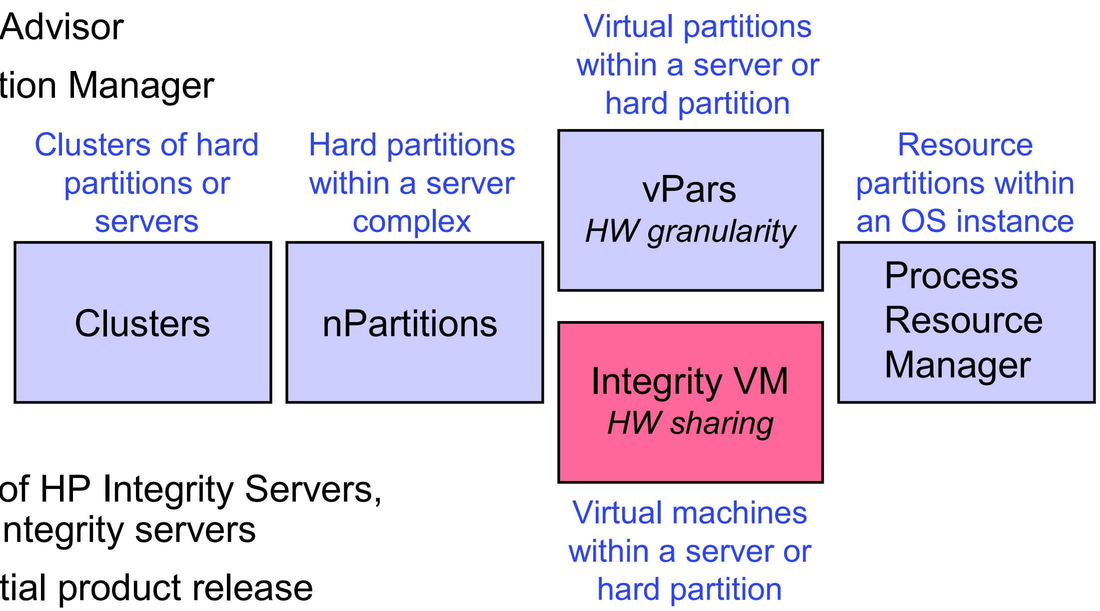
Platform Overview of Virtualization Technology

Surveying the Alternatives (continued)

- **HP**
  - Node Partitions (nPARs) – hardware partitioning solution
  - Virtual Partitions (vPARs) – software solution that only supports HP/UX
  - Virtual Server Environment (VSE) – for HP Integrity and HP 9000 servers
- **Sun**
  - Dynamic Domains – hardware partitioning solution
  - Solaris Containers – all containers (“zones”) share the same copy of Solaris
- **IBM POWER5**
  - LPAR-like solution with I/O and network sharing accomplished via hosted partitions
  - Supports AIX, Linux, and i5/OS

HP Virtual Server Environment (VSE)

- **HP VSE consists of several product components**
  - HP Integrity Essentials Capacity Advisor
  - HP Integrity Essentials Virtualization Manager
  - HP-UX Workload Manager
  - HP Integrity Essentials Global Workload Manager
  - HP Systems Insight Manager
  - HP Integrity Virtual Machines



- **HP Integrity Virtual Machines**
  - Supports low- to high-end range of HP Integrity Servers, Integrity Superdome, and future Integrity servers
  - Support limited to HP-UX with initial product release
  - Supports Uni and SMP guest images with CPU sharing
  - Able to host different guest operating systems
- **"For the most demanding mission-critical workloads, the HP VSE for HP-UX 11i is unmatched"**

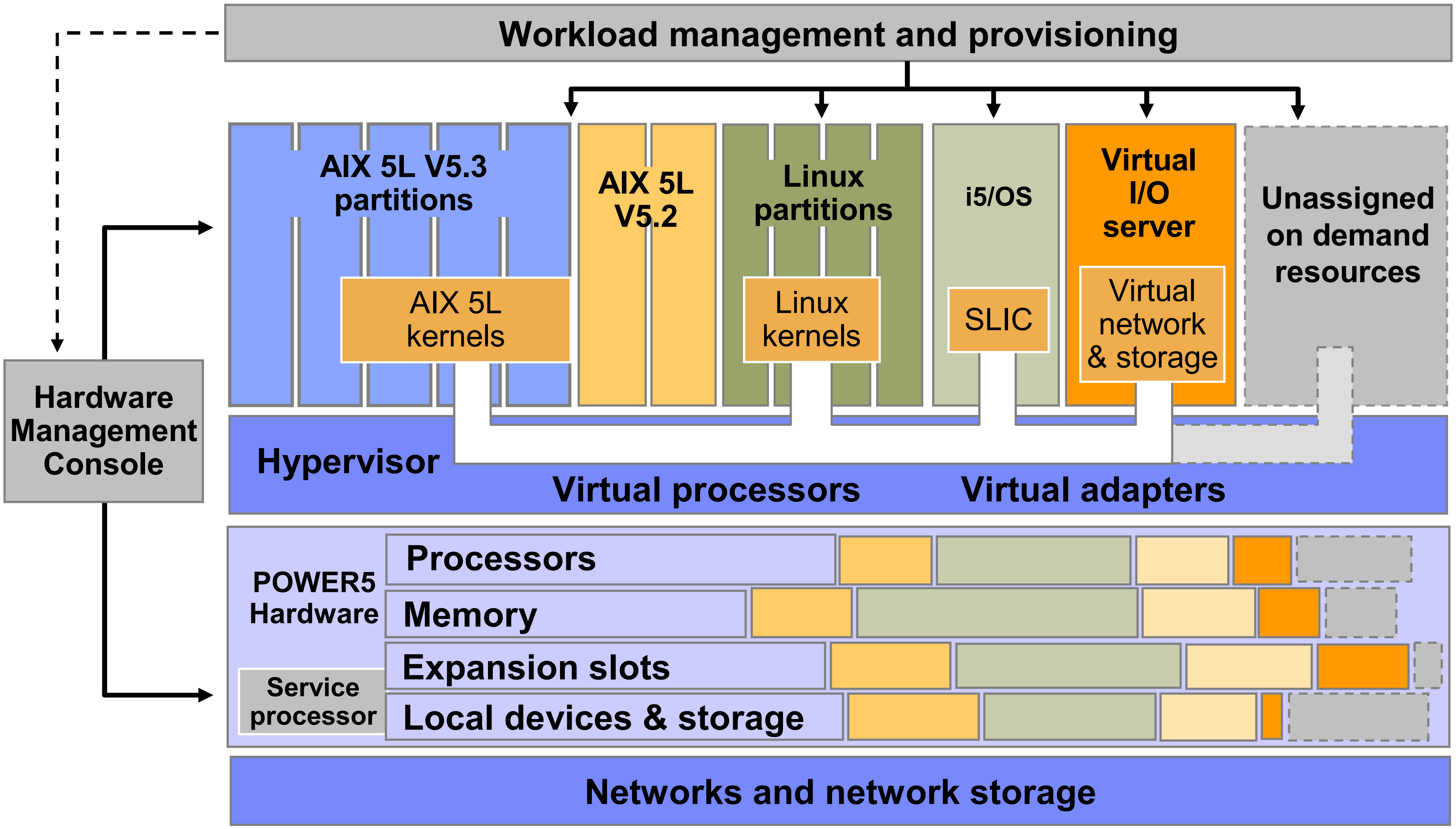
HP Integrity Virtual Machines

- Supports low- to high-end range of HP Integrity Servers, Integrity Superdome, and future Integrity servers
- Support limited to HP-UX with initial product release
- Supports Uni- and SMP-guest images with CPU sharing
- Able to host different guest operating systems

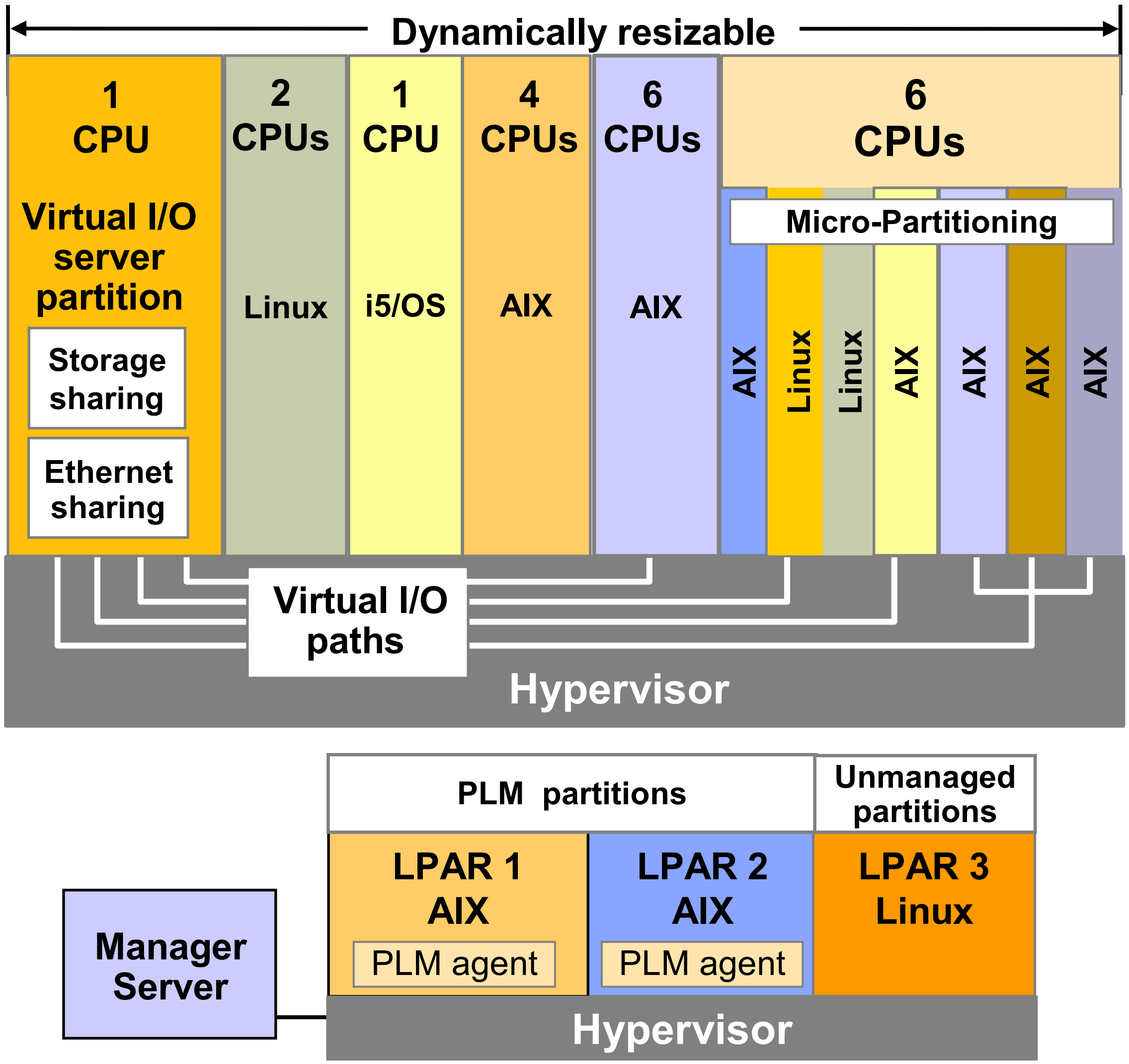
Sun Containers ("Zones")

- **Available with Solaris 10**
- **Similar in concept to z/OS address spaces**
- **Provides ability to host multiple OS instances, not virtual machines**
  - Private name, IP address, and port range
  - Can boot / re-boot zone image
  - Can create a new zone in minutes
  - Separate security, resource management, and failure scopes
  - Private file system options
    - "whole root" – zone owns private, R/W file systems
    - "sparse root" – zone inherits R/O /usr, /sbin, /platform, /lib
- **Instances share the same *Solaris* kernel**
- **Intra-zone networking (memory speeds)**
- **Can share binaries in RAM across zones**

POWER5 Virtualization Architecture



POWER5 Advanced Virtualization Options



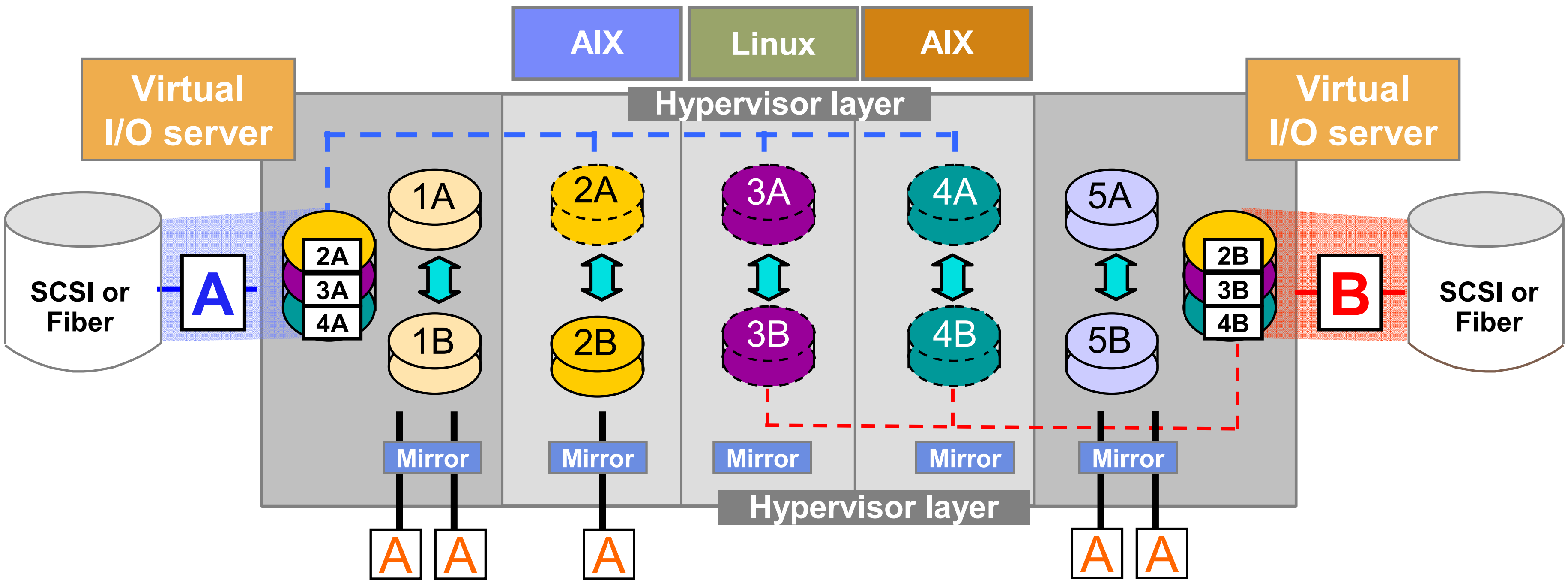
Virtual I/O server
- Shared Ethernet
- Shared SCSI & Fiber Channel attached disk subsystems
- Supports AIX and Linux partitions

Micro-Partitioning
- Share processors across multiple partitions
- Minimum partition size: 1/10th of one CPU
- AIX, Linux, or i5/OS

Partition Load Manager
- Cross-partition workload management for AIX
- Balances processor and memory requests

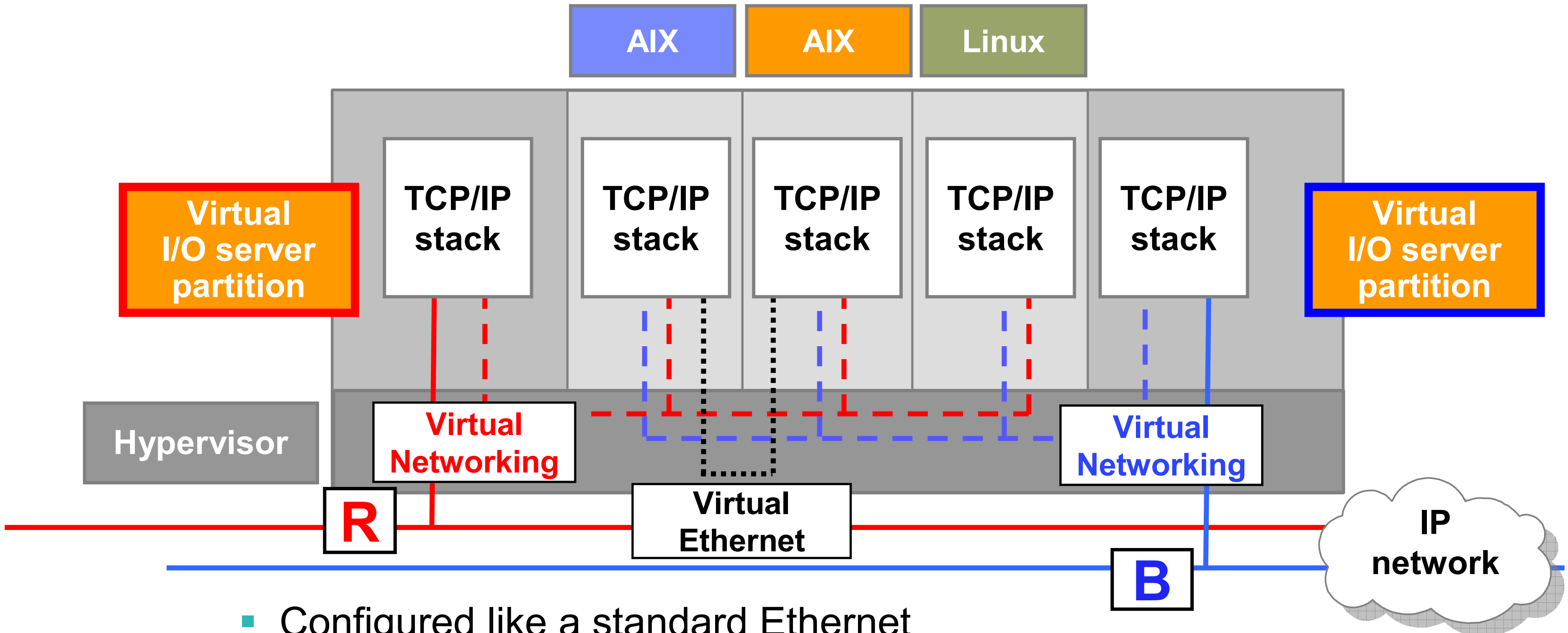
Managed via HMC

POWER5 Virtual I/O Server Disk Sharing



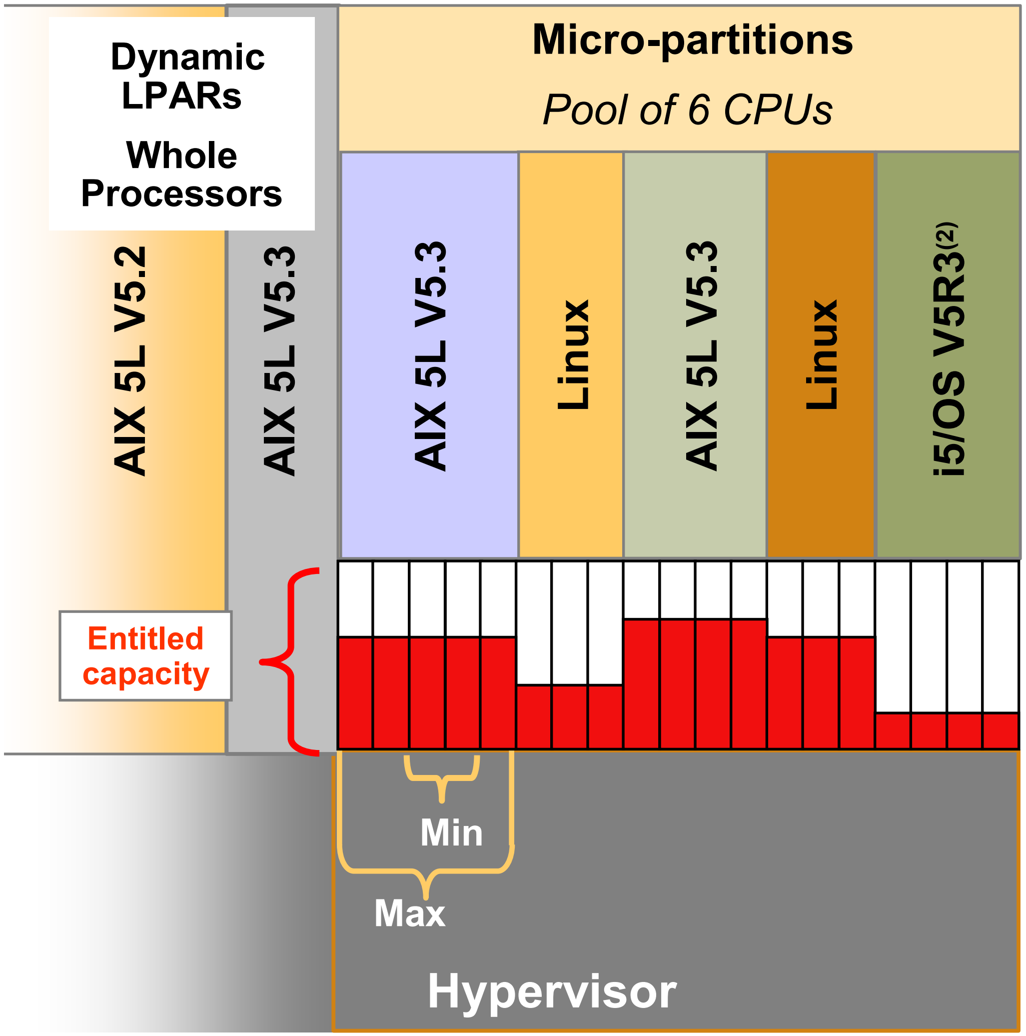
- One physical drive can appear to be multiple logical drives
  - LUNs appear as individual logical drives
  - Hosting partition required for I/O operation
- Minimizes the number of adapters
- Can have mixed configuration (virtual and real adapters)
- SCSI and Fiber supported
- Supports AIX and Linux partitions

POWER5 Virtual I/O Server Ethernet Sharing



- Configured like a standard Ethernet
- IP forwarding / bridging provided by I/O server partition
- Can have multiple connections per partition
- Virtual “MAC” addressing
- Each adapter can support up to 16 virtual Ethernet LANs

POWER5 Micro-Partitioning



Increased number of LPARs
- Micro-Partitions: Up to 254
- Dynamic LPARs: Up to 32

Configured via the HMC

Number of logical processors
- Minimum/maximum

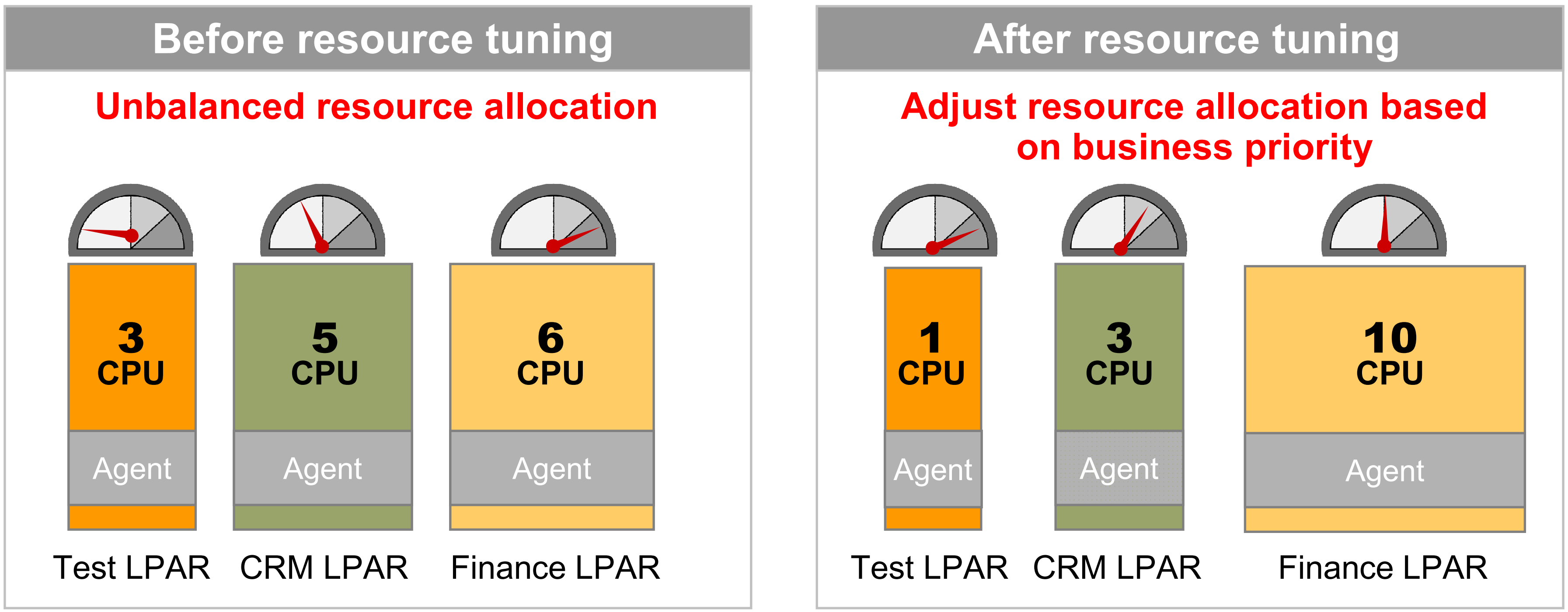
Entitled capacity
- In units of 1/100th of a CPU
- Minimum: 1/10th of a CPU

Variable weight
- % share (priority) of surplus capacity

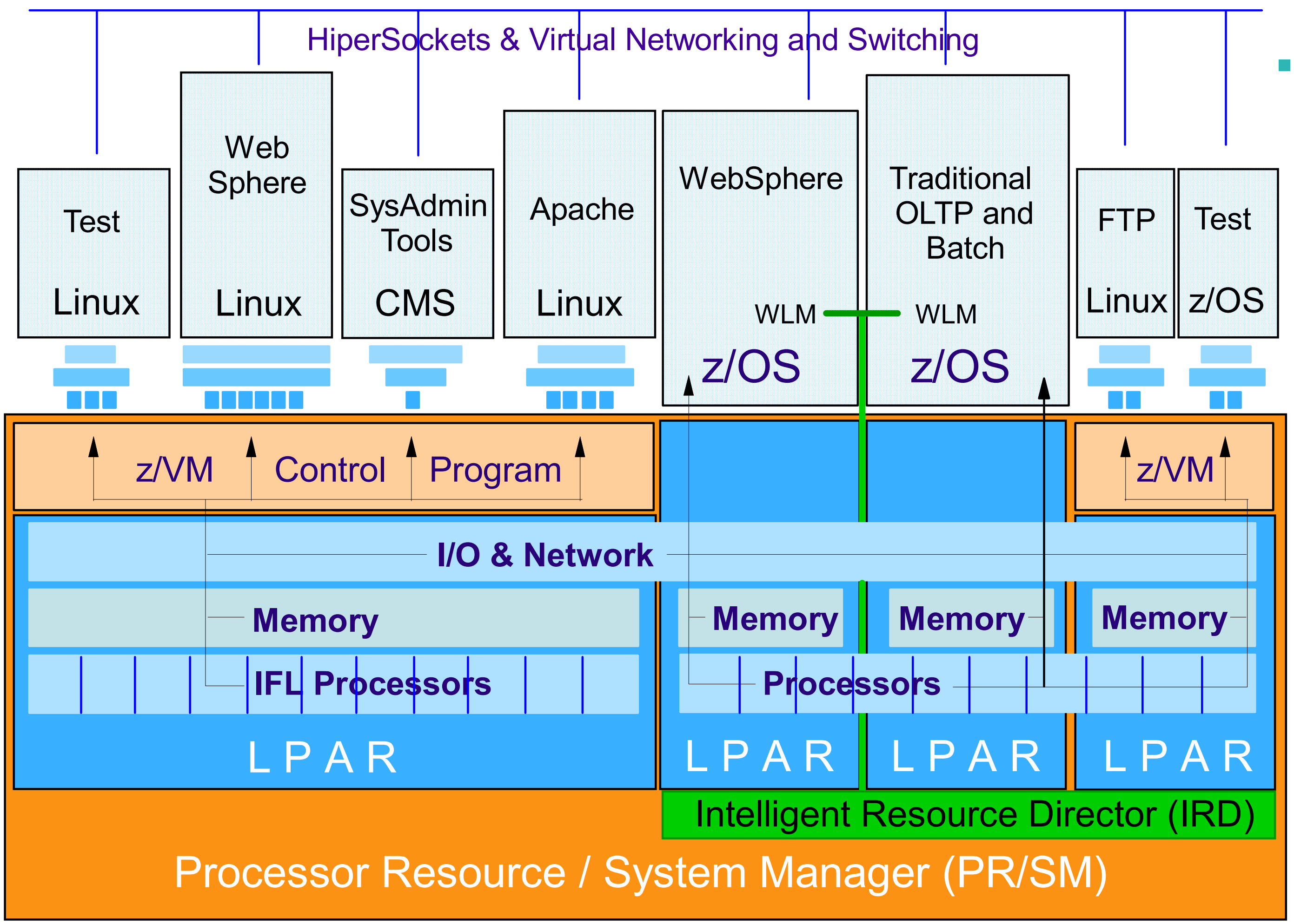
Capped or uncapped partitions

Partition Load Manager for AIX 5L

- **Policy-based, automatic partition resource tuning**
- **Dynamically adjust CPU and memory allocation**



System z9 and zSeries Virtualization Architecture



- **Multi-dimensional virtualization technology**
  - zSeries provides logical (LPAR) and software (z/VM) partitioning
  - PR/SM™ enables highly scalable virtual server hosting for LPAR *and* z/VM virtual machine environments
  - IRD coordinates allocation of CPU and I/O resources among z/OS and non-z/OS LPARs*

* Excluding non-shared resources like Integrated Facility for Linux processors

LPAR and z/VM: World-class Server Virtualization

- **Mainframe Logical Partitioning (LPAR), introduced in 1988, has provided years of business-critical, high-performance server partitioning for the world's largest corporations**
- **z/VM, commercially available since 1972, has supported mixed workloads that require minimal hypervisor overhead, massive scalability, and exceptional levels of availability**
- **Both LPAR and z/VM employ hardware and firmware innovations developed over the years that make virtualization part of the basic componentry of System z9 and zSeries**

System z9 and zSeries Interpretive Execution

Advanced Technology for Virtual Server Hosting

- **Start Interpretive Execution (SIE) instruction**
  - Operand is a state descriptor for an LPAR or virtual machine
  - Accommodates fixed-storage and pageable guests
  - Interception controls allow hypervisor intervention
  - Reduces context switch time
- **zSeries implements two levels of SIE**
  - No performance penalty for running z/VM in an LPAR
  - No shadow page tables required for DAT-on guests
  - Considerable architectural and hardware investment required
    - Potential instruction behavioral differences at each level
    - Multiple control register sets

Additional Mainframe Virtualization Facilities

- **Zone Relocation**
  - SIE capability that provides multiple zero-origin storage regions (logical partitions) on one system
  - Enables I/O subsystem to access partition memory directly, without requiring hypervisor intervention
- **Translation Lookaside Buffers (TLBs)**
  - Large allocation of microprocessor space for TLBs directly benefits virtual server scalability
  - z9-109 and z990 provides a TLB arrangement which advantageously uses two buffers
    - Second-level TLB feeds address translation information to the first-level TLB when the desired virtual address is not contained in the first-level TLB
- **Multiple Image Facility (MIF)**
  - Enables channel sharing among multiple LPARs
  - I/O devices on shared channel paths can be accessed simultaneously by sharing LPARs (or restricted to a subset of sharing LPARs)
- **Logical Channel Subsystems (LCSS) support**
  - Allows a z9-109 and z990 to be configured with up to 1024 channels (512 channels for z890)
  - 256 channels can be configured for each LPAR, with selected channel sharing among LPARs possible

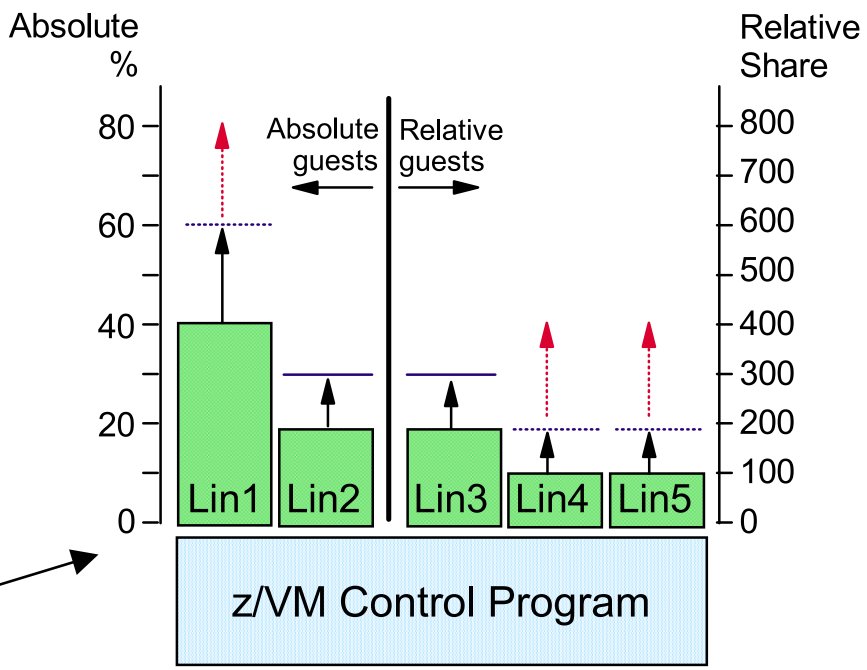
Additional Mainframe Virtualization Facilities

- **I/O Priority Queueing**
  - Allows high-priority workloads to receive preferential access to I/O subsystem
  - Supported by Intelligent Resource Director and virtualized by z/VM
- **HiperSockets**
  - High-speed, security-rich TCP/IP connectivity among LPARs
  - Memory speed communications
- **Adapter Interruption Pass-Through and QDIO Enhanced Buffer-State Management**
  - OSA-Express (Ethernet) and FCP (SCSI) virtual machine I/O can be performed while z/VM guest image is running in SIE mode
  - “Thin” interrupt passed to z/VM Control Program when I/O operation belongs to an idle guest system
- **QDIO Enhanced Buffer-State Management (QEBSM)**
  - Two new machine instructions designed to help eliminate overhead of hypervisor interception
- **Host Page-Management Assist (HPMA)**
  - Interface to z/VM paging and storage management
  - Designed to allow hardware to assign, lock, and unlock page frames without hypervisor assistance
- **Layer 2 (MAC) and Layer 3 (IP) network switching**
  - OSA and z/VM support enables virtual IP and MAC network switching without requiring a hosting partition

z/VM CPU Resource Controls

Highly Granular Sharing of System Resources

- Allocate system resources among virtual machines using SET SHARE command
  - This is a highly flexible and self-managed function of the z/VM Control Program
  - Reserve CPU capacity for peak usage
    - Use it when needed
    - Relinquish the processor cycles for other servers when not needed
  - "Absolute guests" receive top priority
  - The Virtual Machine Resource Manager can be used to monitor and adjust remaining capacity allocated to "Relative guests"



z/VM Directory Entries

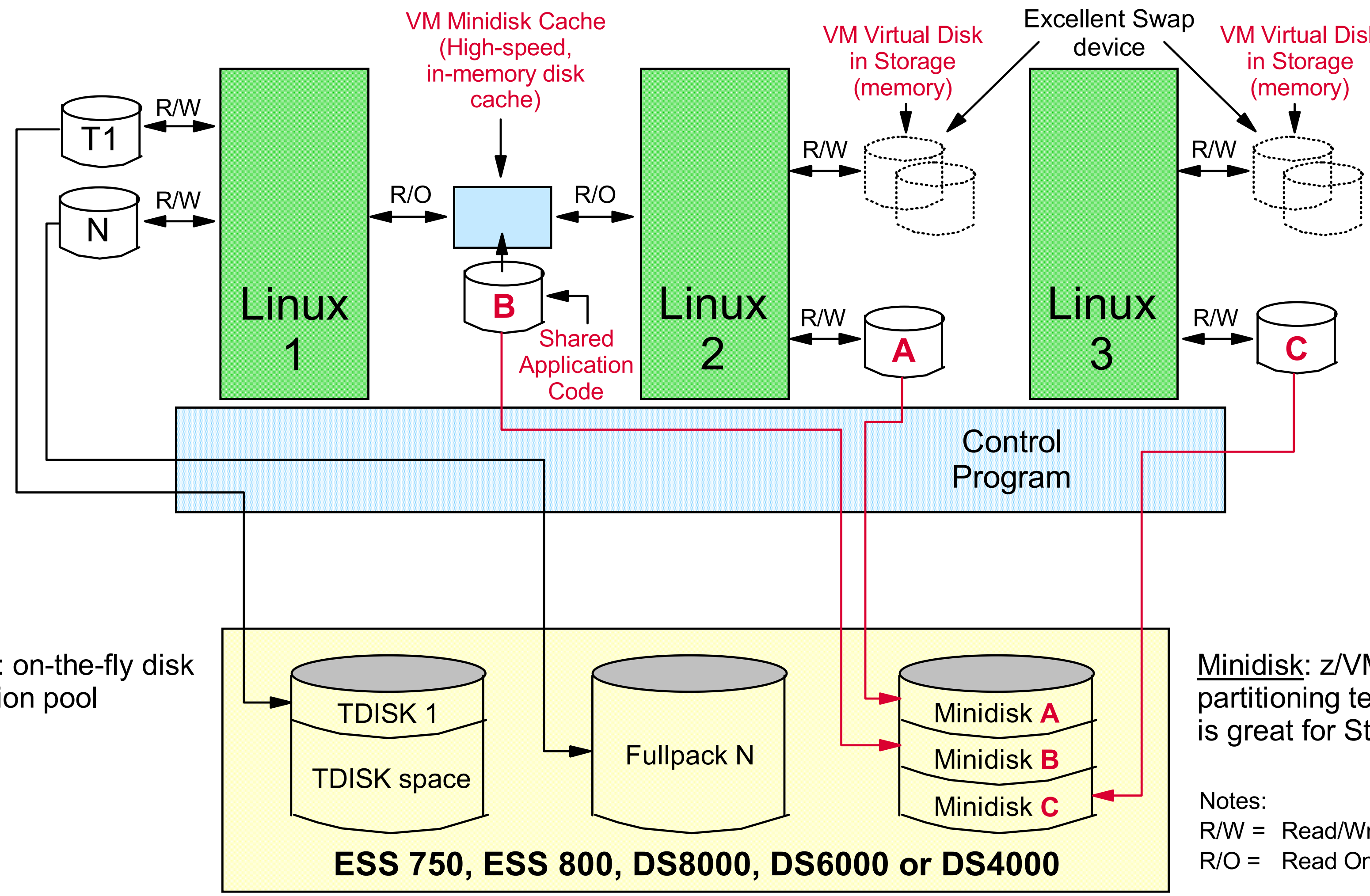
```
SHARE Lin1 ABSOLUTE 40% ABSOLUTE 60% LIMITSOFT
SHARE Lin2 ABSOLUTE 20% ABSOLUTE 30% LIMITHARD

SHARE Lin3 RELATIVE 200 RELATIVE 300 LIMITHARD
SHARE Lin4 RELATIVE 100 RELATIVE 200 LIMITSOFT
SHARE Lin5 RELATIVE 100 RELATIVE 200 LIMITSOFT
```

Notes:

(dotted line) = limit can be exceeded if unused capacity is available (limitsoft)

(solid line) = limit will not be exceeded (limithard)

z/VM Disk Virtualization Technology



TDISK: on-the-fly disk allocation pool

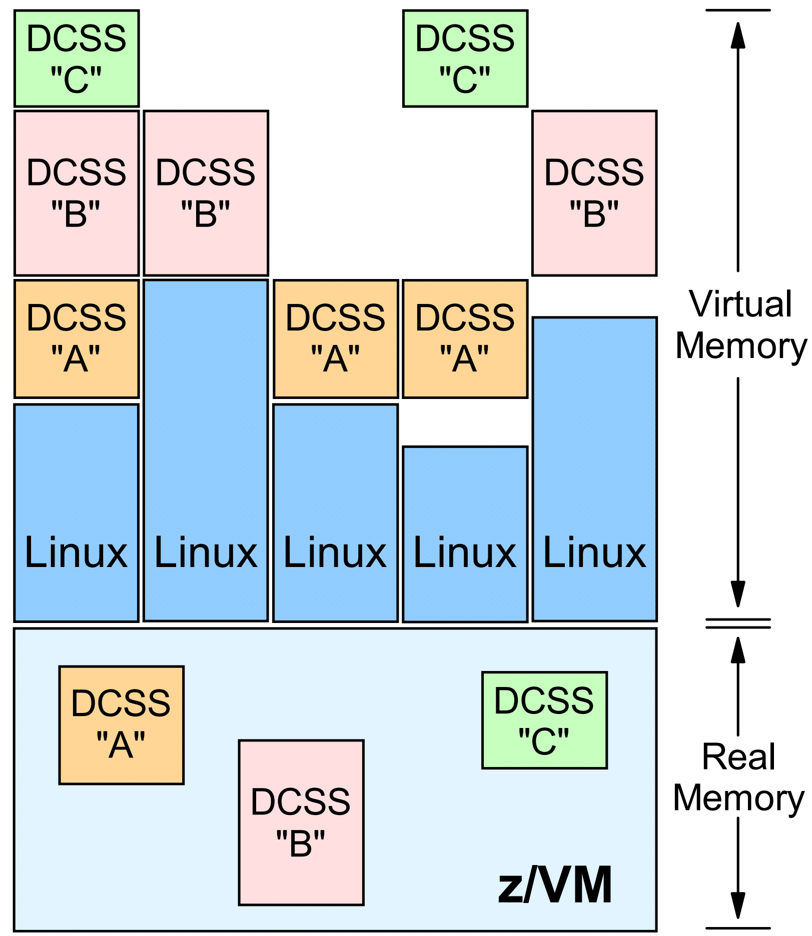
Minidisk: z/VM disk partitioning technology is great for Staging!

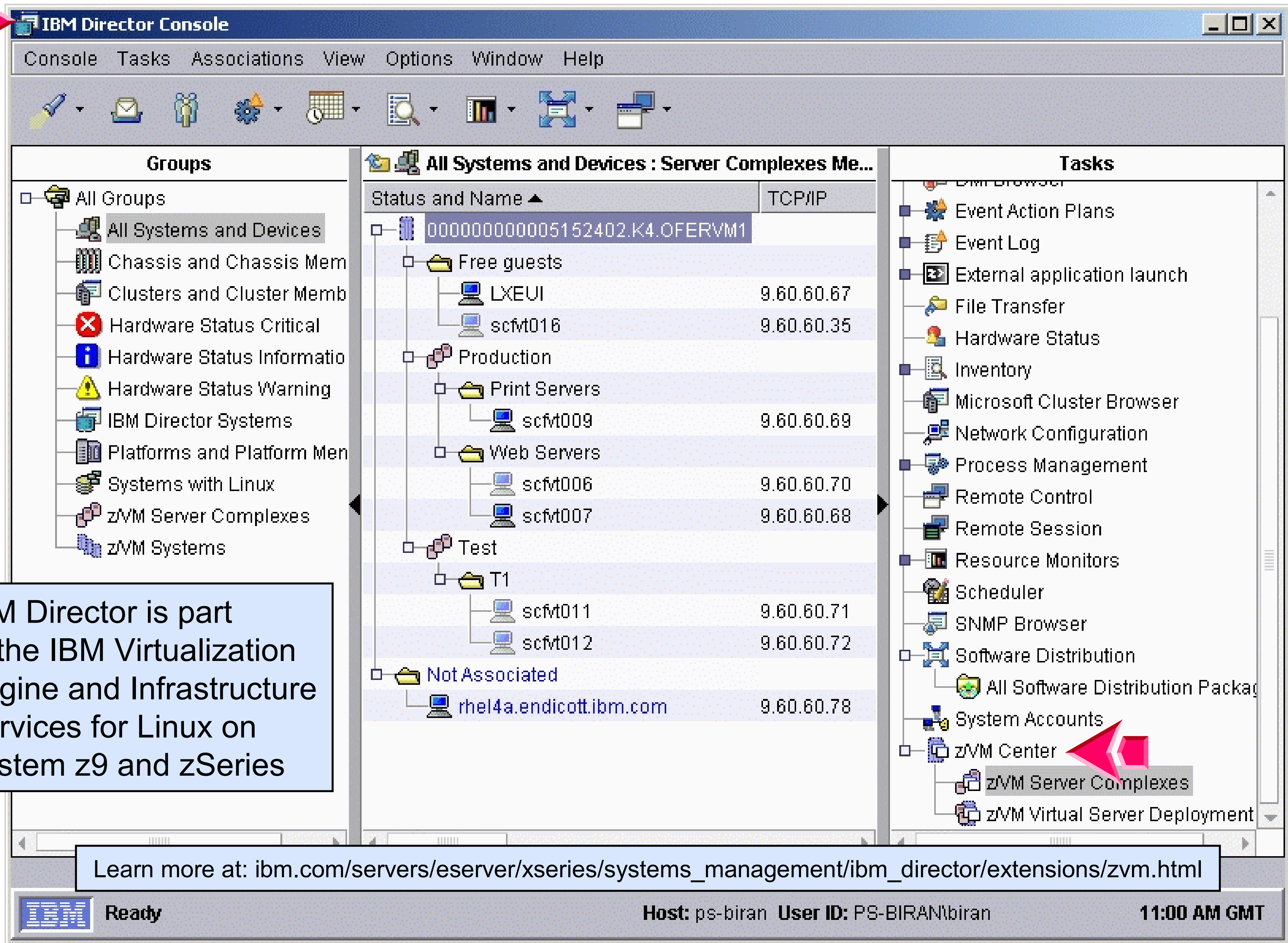
Notes:
R/W = Read/Write
R/O = Read Only

Linux and z/VM Technology Exploitation

Linux Exploitation of z/VM Discontiguous Saved Segments (DCSS)

- DCSS support is a z/VM exclusive
  - Share a single, real memory area among multiple virtual machines
  - High-performance data access
  - Can reduce real memory utilization
- Linux exploitation support available today
  - Execute-in-place (xip2) file system
  - DCSS memory locations can reside outside the defined virtual machine configuration
  - Access to file system is at memory speeds; executables are invoked directly out of the file system (no data movement required)
  - Avoids duplication of virtual memory and data stored on disks
  - Enables throughput benefits for Linux guest images and enhances overall system performance and scalability

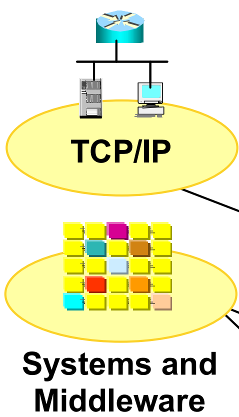




IBM Director is part of the IBM Virtualization Engine and Infrastructure Services for Linux on System z9 and zSeries

Learn more at: ibm.com/servers/eserver/xseries/systems_management/ibm_director/extensions/zvm.html

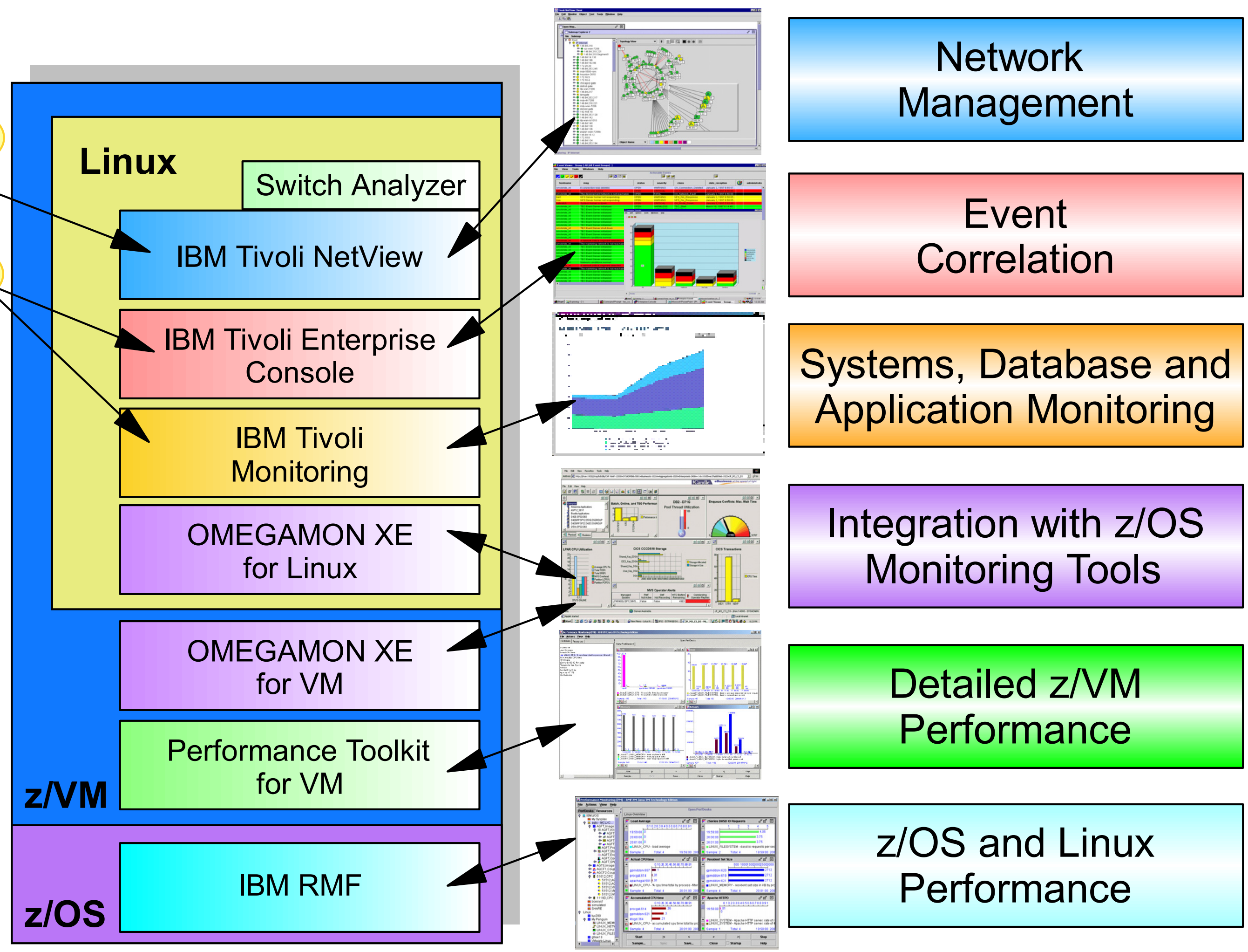
Monitoring Linux on System z and zSeries



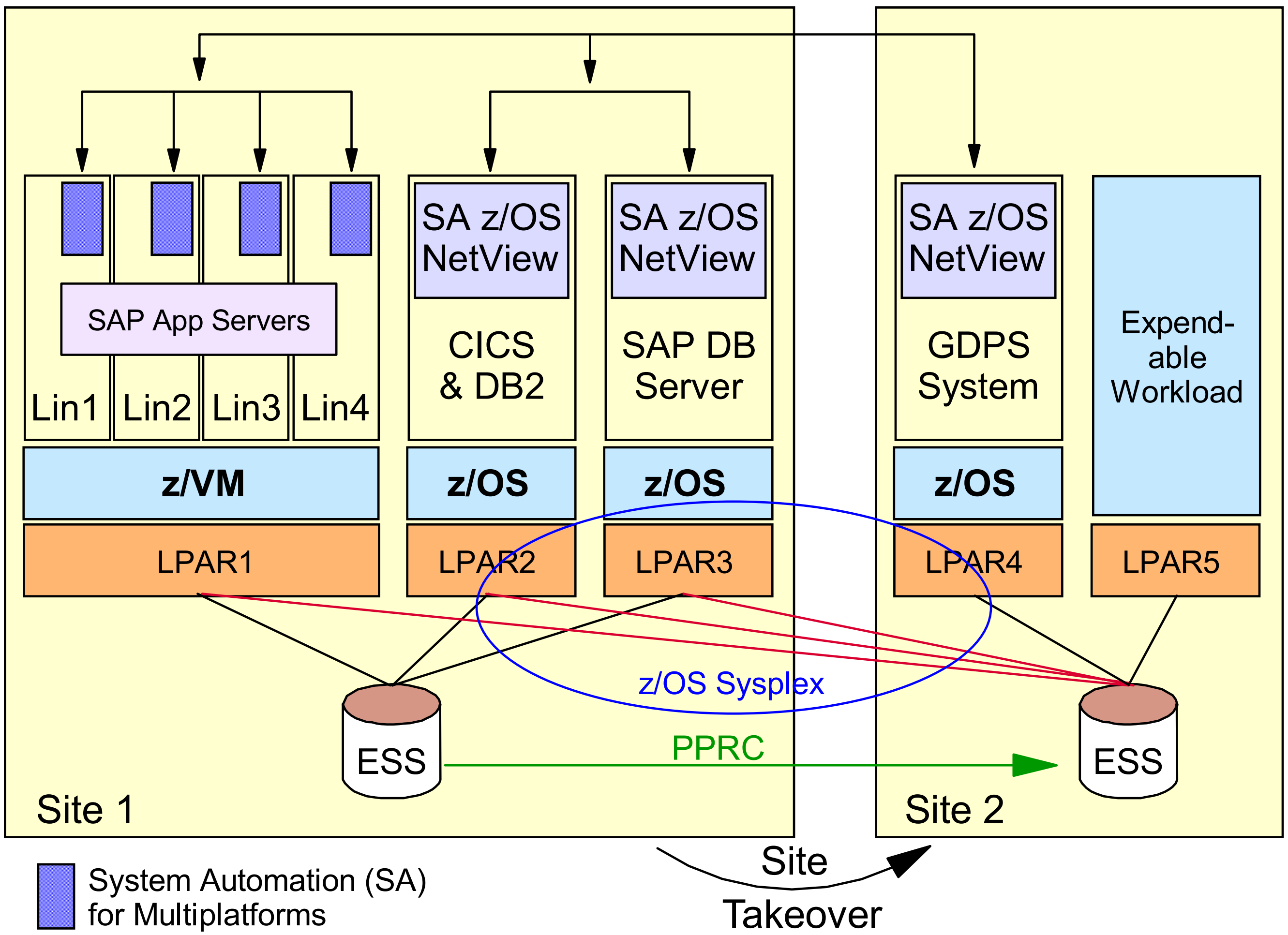
Measure and manage guest systems from the "outside" using z/VM tools and facilities

Complement the environment with Linux-hosted Tivoli (and ISV) solutions

Optionally integrate with a z/OS infrastructure



GDPS/PPRC Multiplatform Resiliency for zSeries



- Designed for customers with distributed applications
- SAP application server running on Linux for zSeries
- SAP DB server running on z/OS
- Coordinated near-continuous availability and DR solution for z/OS, Linux guests, and z/VM
- Uses z/VM HyperSwap function to switch to secondary disks
- Sysplex support allows for site recovery

IBM System z Virtualization Differentiation

A "Genetic" View of Virtualization Technologies

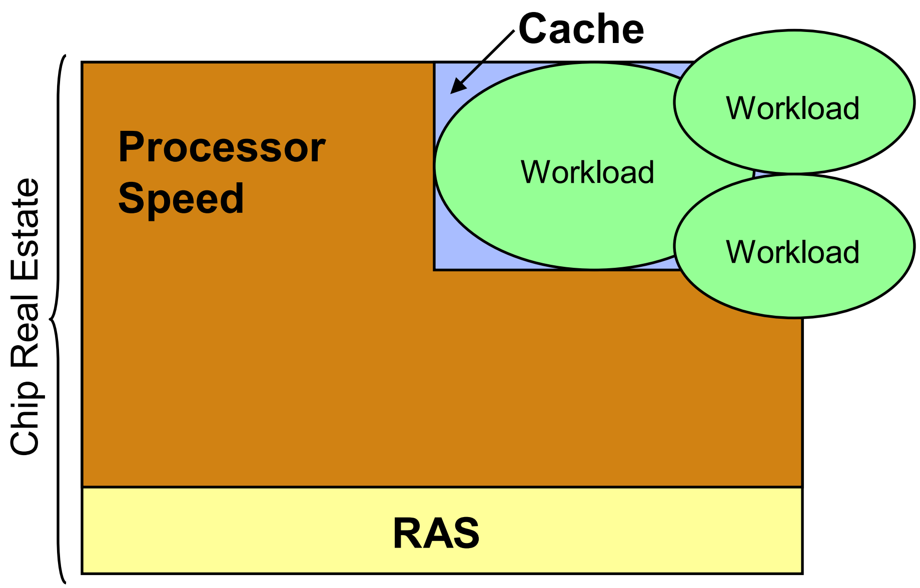
- **Today's server virtualization capabilities are influenced by the heritage of the platform**
  - System z
    - Mainframes have hosted mixed workloads and have optimized resource sharing ever since their introduction in 1964
    - Mainframe investments have been driven by the business community, where high qualities of service are essential
  - System p
    - Unix systems achieved early success in the scientific and engineering community, where high-performance number crunching was a priority
    - Rebooting a failed server was not considered a serious problem
  - xSeries
    - Intel servers gained popularity as an inexpensive platform offering "application freedom" from the IT organization
    - One application per Intel server was the norm; "Ctl-Alt-Del" was the standard method of solving problems

Implications of Virtualization Genetics

- **Virtualization technology and "awareness" is built into mainframe architectures, hardware, firmware, hypervisors, and operating systems**
  - The technology is very responsive and stable, "hardened" by 40 years of usage, feedback, and iterative development
- **Virtualization technology is largely an "add-on" for pSeries and Intel system designs**
  - Longer feedback paths reduce responsiveness and stability
  - The technology does not enjoy years of high-stress deployment in a wide-range of client environments
- **Nonetheless...**
  - One server platform does not satisfy all client requirements
  - A multi-platform deployment of virtualization technology offers maximum responsiveness and efficient resource utilization for on demand computing
  - Exploiting System z9 and zSeries in this environment is good business for mainframe clients
  - Tivoli solutions and the Virtualization Engine can help clients manage an environment that has a mix of server technologies

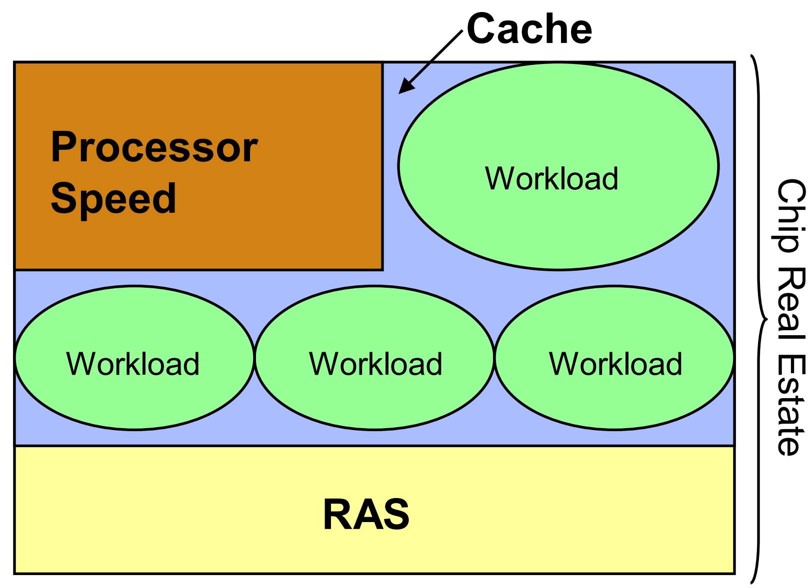
Chip Design Affects Virtualization Capabilities

Replicated Server Chip Design



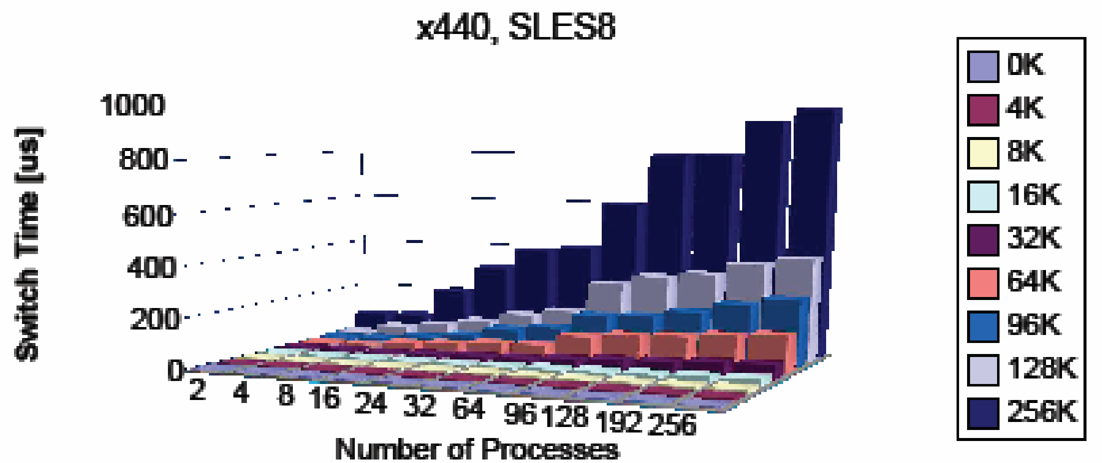
- Mixed workloads stress cache usage, requiring more context switches
- Working sets may be too large to fit in cache
- “Fast” processor speed is not fully realized due to cache misses

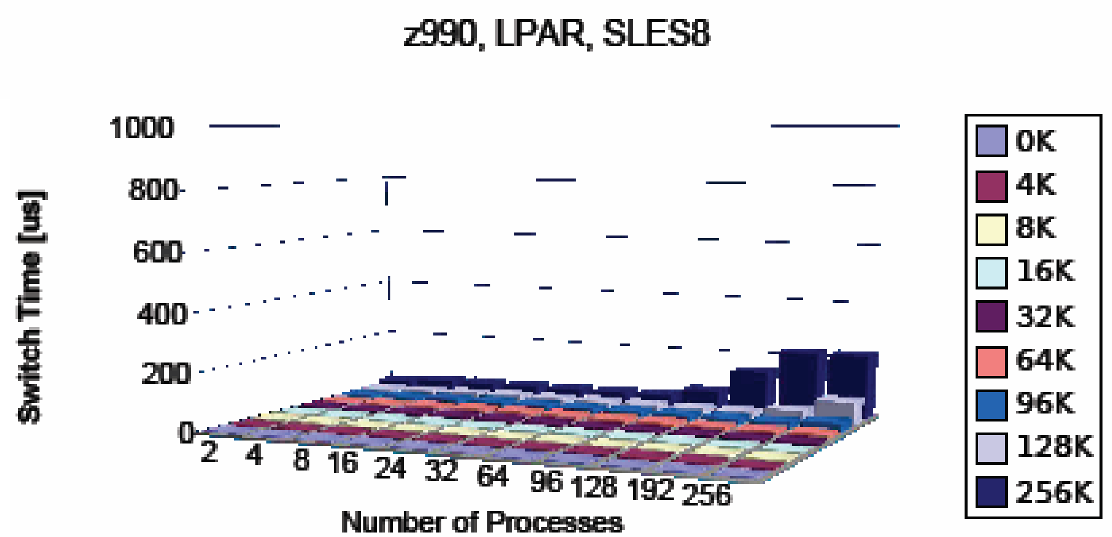
Consolidated Server Chip Design



- zSeries cache is able to contain more working sets
- Processor speed is optimized by increased cache usage
- Additional RAS function is beneficial for mixed workloads

Scalability Considerations: Context Switching





- Time required to perform context switching is an indication of memory time
- Virtualization, by definition, involves context switching

System z Partitioning Leadership Support

- **Highest levels of RAS built into the hardware**
- **Non-disruptive On/Off Capacity on Demand**
- **Linux and z/OS application integration**
- **Highly granular allocation of hardware assets**
  - Add “small” server images to existing config with minimal impact to other server images
- **Large-scale server hosting**
  - Potentially hundreds of server images
- **Resource consumption recording / reporting**
  - Capture data at hypervisor level (CP Monitor)
  - Useful for charge-back, capacity planning, problem determination, and fix verification
- **Hot stand-by without the hardware expense**
  - Idle backup images ready to run (or be booted) if primary servers fail
- **Autonomic, non-disruptive disk failover to secondary storage subsystem**
- **Architecture simulation**
  - Satisfy server configuration requirements without suffering expense of real hardware
- **In-memory application sharing**
  - Share program executables among multiple server images
- **Server-memory-cached disk I/O**
  - High-speed read access to files on disk
- **Virtual Disks in Storage**
  - High-speed read and write access to files in memory (excellent swap devices for Linux)
- **Built-in console message routing**
  - Route messages from all virtual servers to a single virtual machine (system automation)
- **Virtual Machine Resource Manager**
- **“Hands free” auto-logon of server images**
  - Using z/VM “Autolog” support
- **Initiate operating system shutdown from “outside” the server image**
  - Without requiring agent running on guest operating system
- **Up to 256 Linux servers can share a single zSeries cryptographic card using z/VM**
- **Clone, patch, and “go live” without outage and with easy rollback**

System z Versus POWER5 Considerations

- **More granular allocation of CPU resources with z/VM**
- **POWER5 shared I/O and networking requires hosting partition(s)**
- **z/VM shared memory model allows sharing and dynamic allocation of real memory to virtual server images**
  - Innovations like DCSS support exploit the shared memory model for enhanced scalability and operational simplification
- **Potentially faster virtual server creation / provisioning with z/VM**
- **System z9 and zSeries offers a more mature LPAR environment**
- **Highly compute-intensive workload is a better fit for POWER5**
- **Customer acceptance**
  - Linux on zSeries has a 5-year history of success
  - Linux on Power is relatively new

System z Versus VMware Considerations

- **IBM mainframes enjoy significant single-system scalability advantages over VMware systems**
  - CPU utilization and granularity
  - Sophisticated paging subsystem
  - Firmware-assisted I/O, networking, shared memory management
- **Some Linux workloads may be constrained when running on VMware**
  - VMware ESX Server 3 allows no more than 4 CPUs per guest image (and only if 4 real CPUs exist on the Intel server)…z/VM V5 supports up to 64 CPUs per virtual machine and 24 real CPUs (with plans to go beyond 24 real CPUs)
  - VMware ESX Server 3 limits virtual machines to 16 GB of memory…z/VM V5.2 supports virtual machines with 128 GB of memory (or more)
  - I/O-intensive workloads can suffer considerable overhead with VMware
    - Timely interrupt processing can be a problem for VMware
    - Mainframe I/O bandwidth is much larger than Intel/AMD-based systems
  - VMware ESX support for 64-bit guests is “experimental”…z/VM can support a mix of 32-bit and 64-bit guest images today
- **Hosting multiple z/VM images on a single server (via LPAR) can enhance failover options, I/O sharing, and workload distribution**
- **VMware guests cannot add or remove resources without being rebooted**

System z Versus VMware Considerations

- **The current generation of Intel and AMD processors are not designed to support virtualization**
  - But…hardware virtualization support is coming (IVT and Pacifica)
- **In general, VMware users are advised to be careful over-committing virtual memory to real memory**
- **VMware VirtualCenter does not support Enterprise Workload Manager (EWLM) for end-to-end resource allocation across a datacenter**

"The advantages of virtualization in both management and efficiency are major contributors to the renaissance of the IBM zSeries, which has a richer history and greater intellectual property associated with virtualization than does any other platform."

- Gordon Haff, Illuminata, "IBM Fits Linux to Power", 9/13/2004

When Do You Need More than "Good Enough"?

Making the Case for Mainframe Virtualization

- **When workload growth and decline is difficult to predict (be it production, development, or test & assurance systems)**
- **When business results suffer as a result of IT resources not matching customer demand**
- **When your IT staff wants to optimize their productivity for deploying and managing virtual servers**
- **When innovation is stifled because your staff cannot experiment or develop new solutions using existing resources**
- **When speed to market affects your business results**
- **When your server applications need fast and flexible access to z/OS data and applications**
- **When business resiliency is a high priority**
- **When you want more control over your environmental expenses (e.g., floor space, cooling)**

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DB2, e-business logo, ESCON, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/390, System z9, VM/ESA, VSE/ESA, WebSphere, xSeries, z/OS, zSeries, z/VM.

The following are trademarks or registered trademarks of other companies

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries.
LINUX is a registered trademark of Linux Torvalds in the United States and other countries.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.
Intel is a registered trademark of Intel Corporation.
* All other products may be trademarks or registered trademarks of their respective companies.

NOTES:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

RETURN TO INDEX