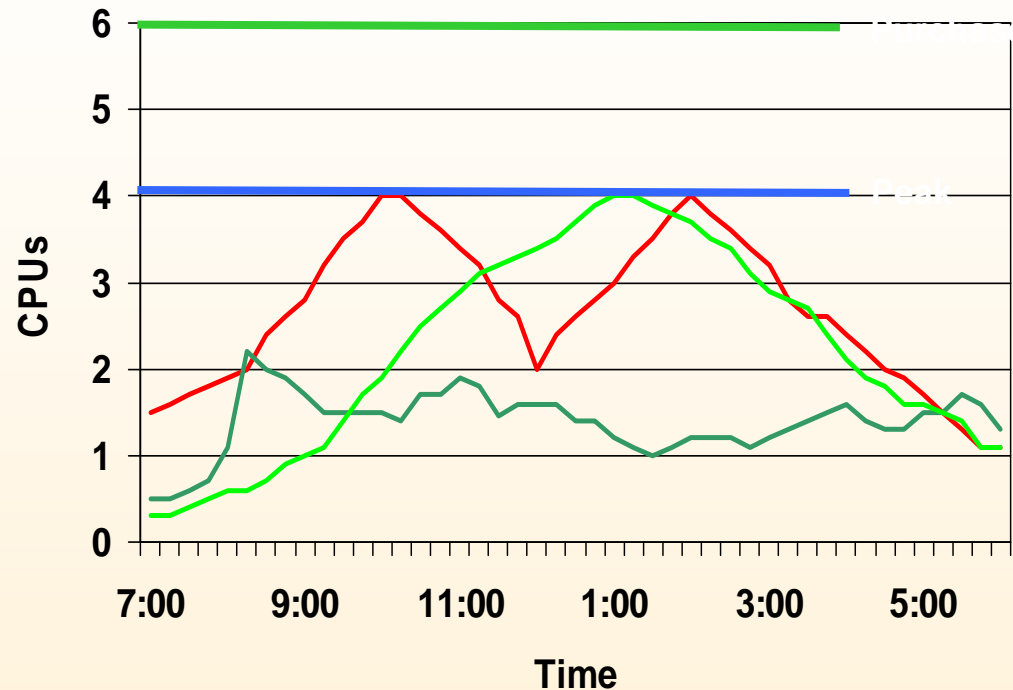


IBM Power Systems p AIX/UNIX/Linux Virtualization Basics

**Key Terminology:
Practical Implementations
of Virtualization
& Live Partition Mobility**

Server Workload Considerations

- Traditional Server Capacity
 - Workloads fluctuate differently throughout the day and year
 - Workloads can fluctuate greatly moment to moment
 - Servers are purchased to handle individual unknown peaks
 - Unused resources can not be moved between servers



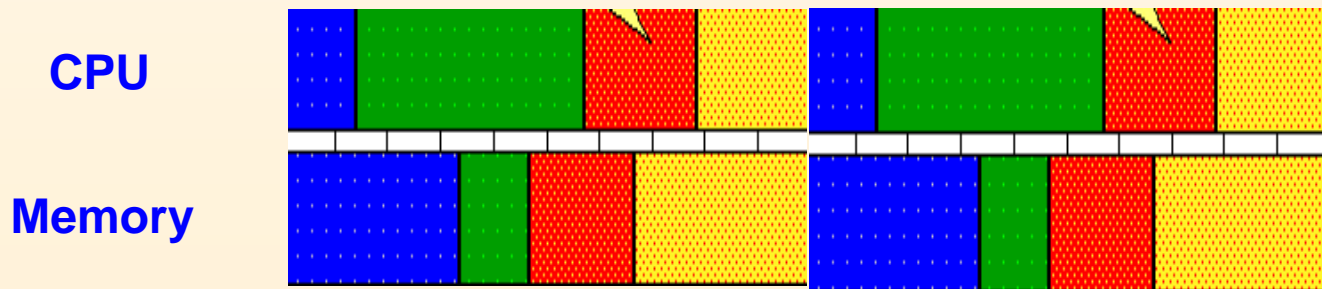
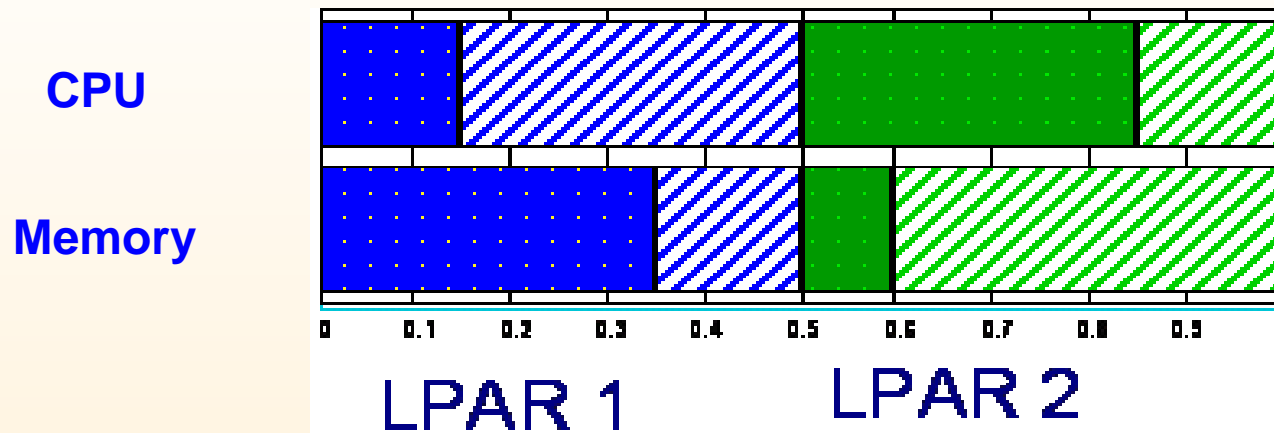


LPAR (Logical Partitions) – Separate O/S Images

AIX Workload Manager (WLM)

Multiple Applications (processes) controlled (fenced off) with a single AIX image

- Improved administrative efficiency by reducing the number of AIX images to maintain



AIX Workload Partitions

Separate regions of application space within a single AIX image

- Improved administrative efficiency by reducing the number of AIX images to maintain

- **Software partitioned system capacity**

- Each Workload Partition obtains a regulated share of system resources
- Each Workload Partition can have unique network, filesystems and security

- **Two types of Workload Partitions**

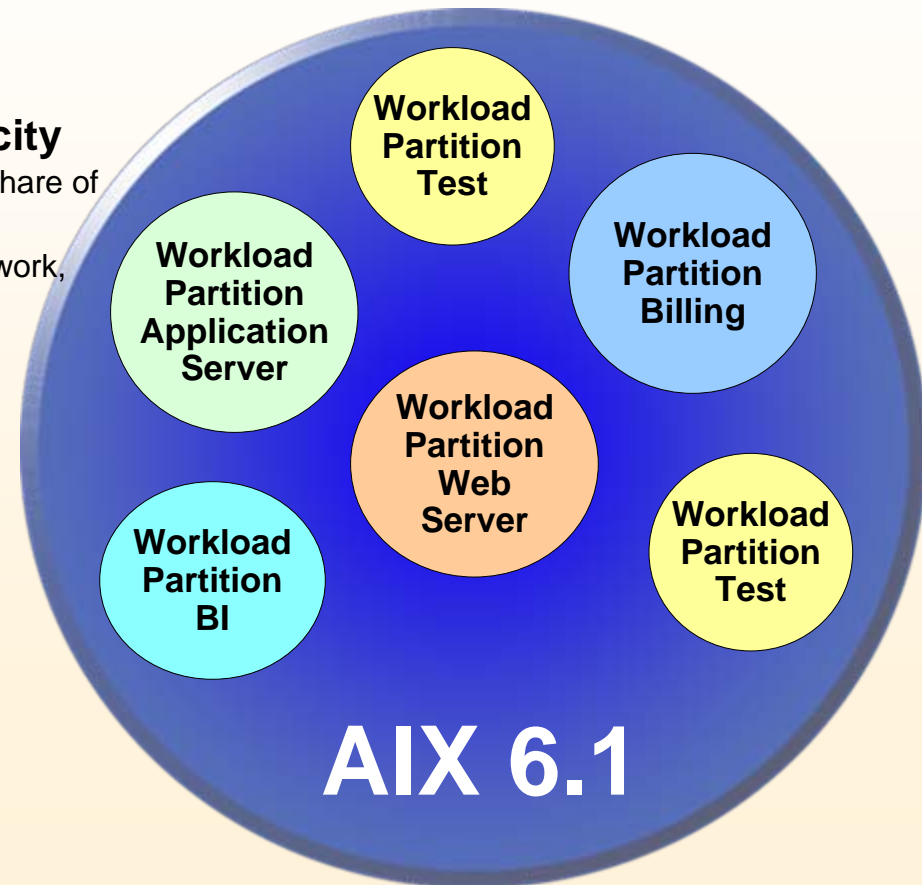
- System Partitions
- Application Partitions

- **Separate administrative control**

- Each System Workload partition is a separate administrative and security domain

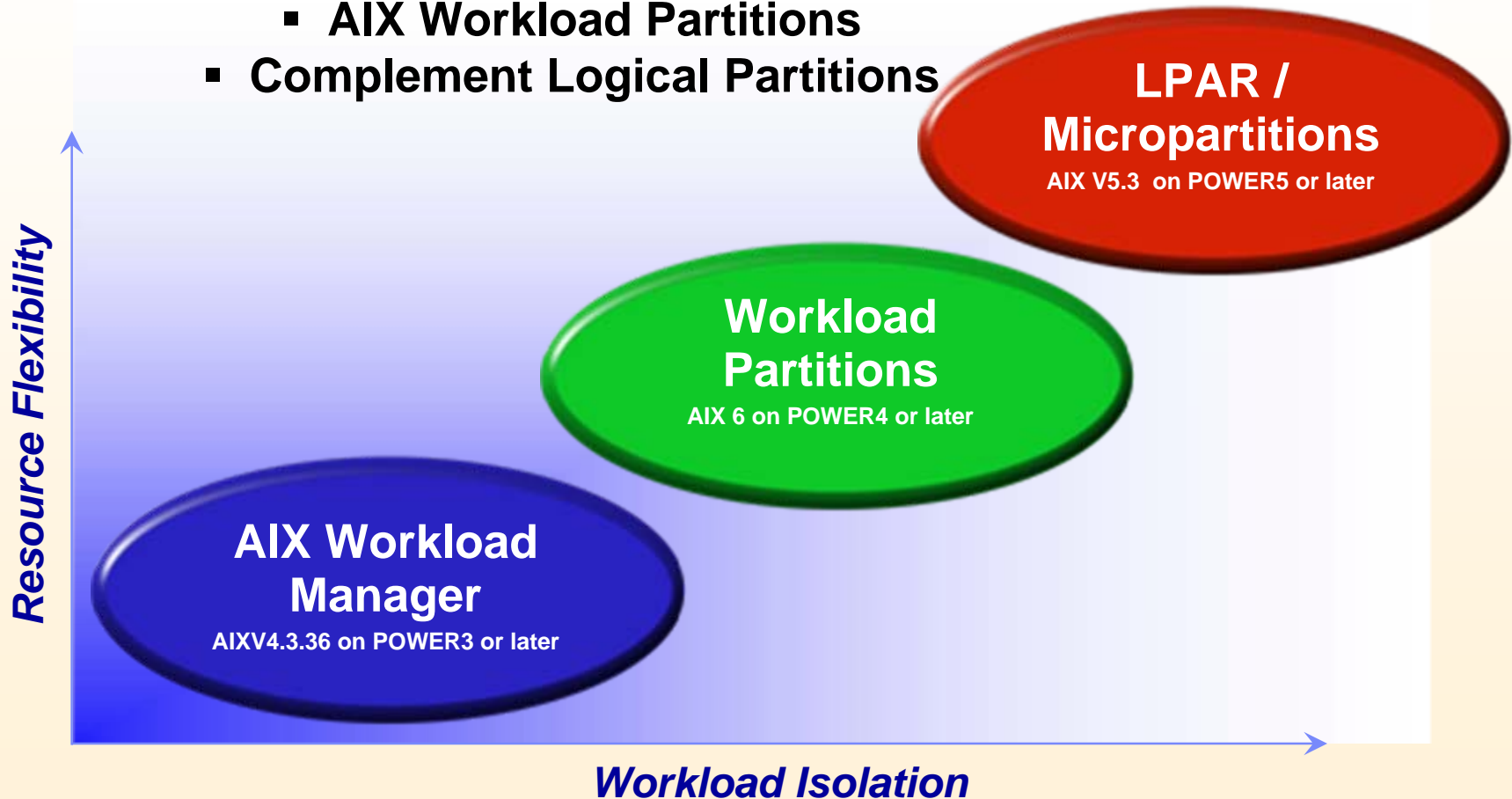
- **Shared system resources**

Operating System, I/O, Processor, Memory



PowerVM AIX Virtualization Continuum

- AIX Workload Partitions
- Complement Logical Partitions



Two WPAR AIX Offerings...

- **AIX 6**

- Workload Partitions (WPAR) included in AIX 6
- Element (single system) WPAR Management



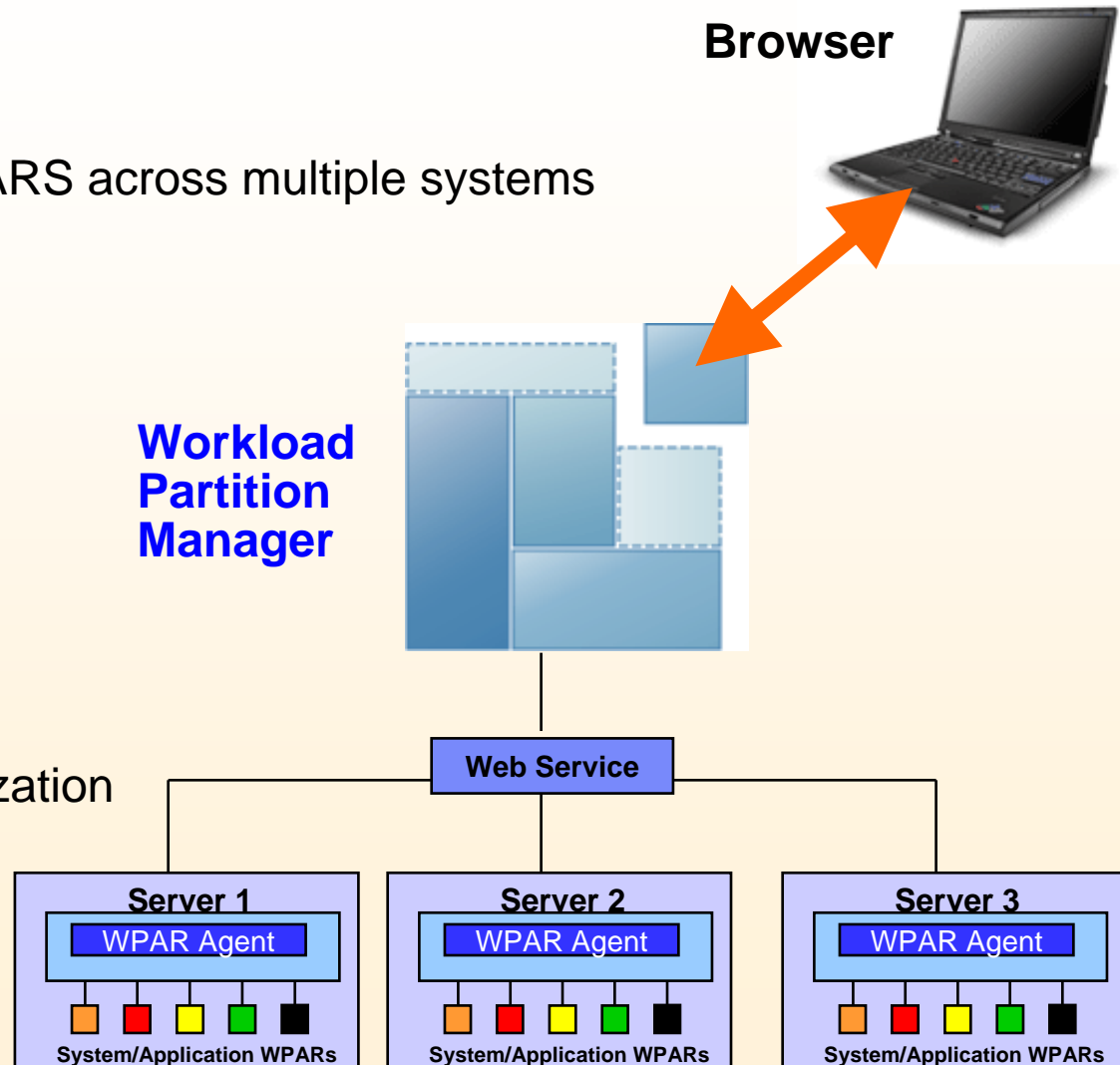
- **Workload Partitions Manager™**

- Enablement for Live Application Mobility
- Cross System Management for Workload Partitions
- Automated, Policy-based Application Mobility
- Part of IBM System Director Family



Workload Partitions Manager

- Management of WPARS across multiple systems
- Lifecycle operations
- Single Console for:
 - Graphical Interface
 - Create & Remove
 - Start & Stop
 - Checkpoint & Restart
 - Monitoring & Reporting
 - Manual Relocation
 - Automated Relocation
 - Policy Driven Change
- Infrastructure Optimization
- Load Balancing



Graphical WPAR Manager & Application Mobility

The screenshot shows the Workload Partition Manager web interface. The browser window title is "Workload Partition Manager - Mozilla Firefox". The page header includes "Workload Partition Manager Welcome wparmgrID" and "Help | Logout". The left sidebar contains a navigation menu with the following sections:

- Guided Activities
 - Create Workload Partition
 - Create Migration Rules
 - Create Stack
 - Migrate Workload Partitions
- Systems
 - Managed Servers
 - Workload Partitions
- Monitoring
 - Event Activity
 - Event History
 - Performance
- Mobility
 - Migration Rules
 - Target Ranking
 - Stacks/Tiers
- Troubleshooting
 - Log and Trace
- WPAR Manager Settings
 - Configuration

The main content area is titled "Managed servers" and includes the text: "Managed servers are systems controlled by the Workload Partition Manager." Below this text are buttons for "Remove", "Show Workload Partitions" (with a dropdown arrow), and "Go". There are also icons for a folder and a document. The main content area displays a table of managed servers:

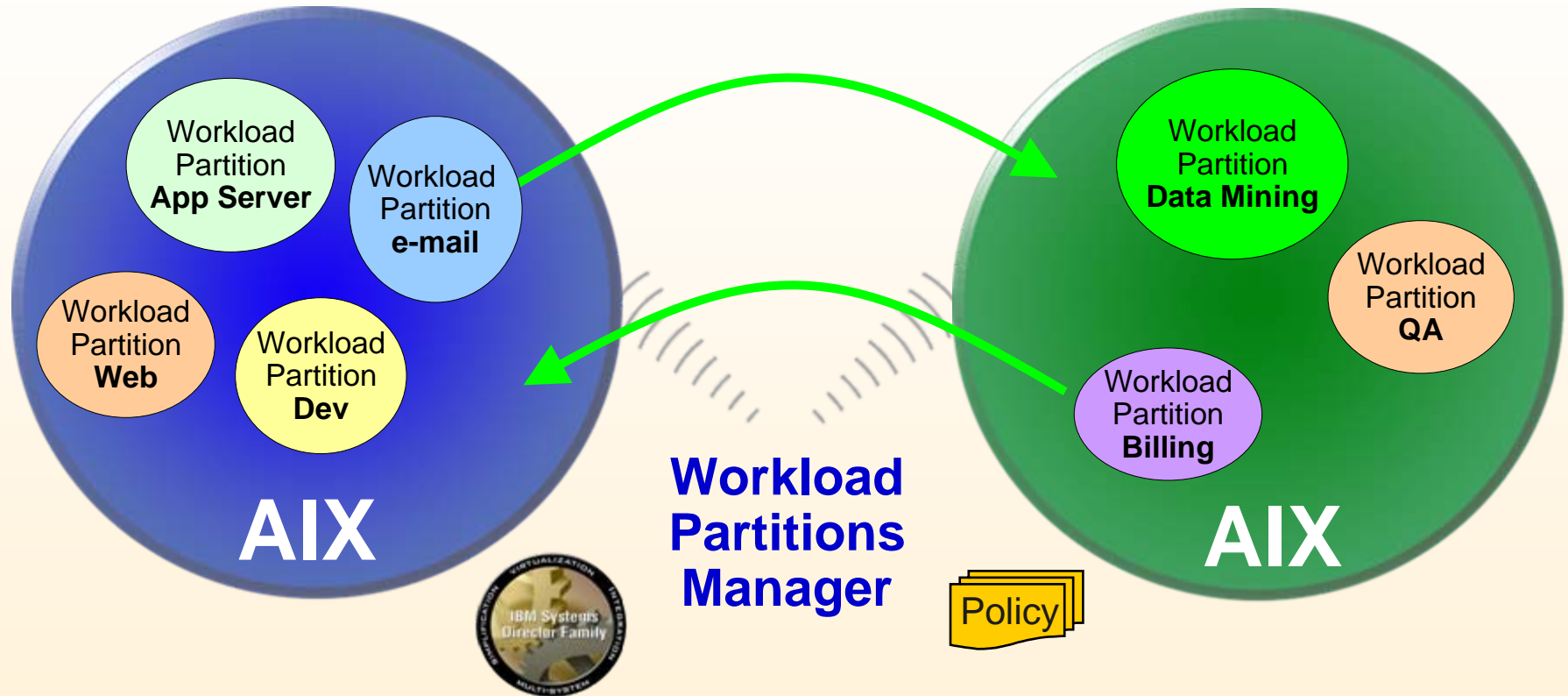
Select	Id	Hostname	State	Description	IP address	# Workload Partitions
<input type="checkbox"/>	5	server5	■	this is server5	10.10.10.5	0
<input type="checkbox"/>	4	server4	■	this is server4	10.10.10.4	0
<input type="checkbox"/>	3	server3	■	this is server3	10.10.10.3	0
<input type="checkbox"/>	2	server2	■	this is server2	10.10.10.2	0
<input type="checkbox"/>	1	server1	■	this is server1	10.10.10.1	1
<input type="checkbox"/>	0	-	■	Unassigned	-	0

At the bottom of the table, there is a pagination control showing "Page 1 of 1".

**Workload
Partition
Manager**

AIX Live Application Mobility

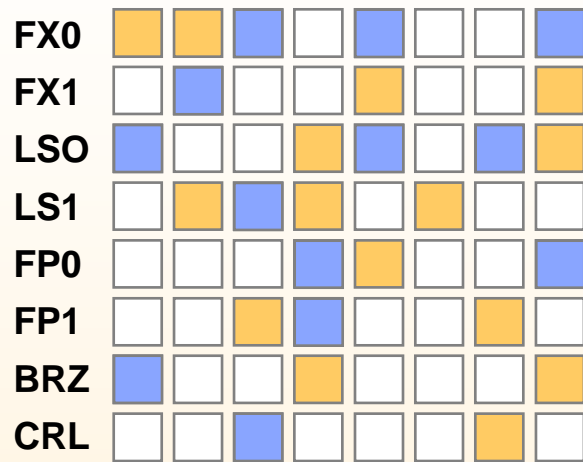
Move a running Workload Partition from one server to another for outage avoidance and multi-system workload balancing



Works on any hardware supported by AIX 6, including POWER5 and POWER4

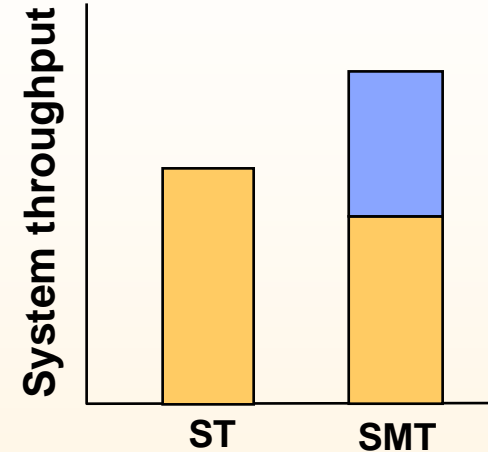
IBM System Simultaneous multithreading

POWER5 & POWER6 (simultaneous multithreading)



■ Thread0 active
 □ No thread active
 ■ Thread1 active

Appears as four CPUs per chip to the operating system (AIX 5L V5.3 and Linux)

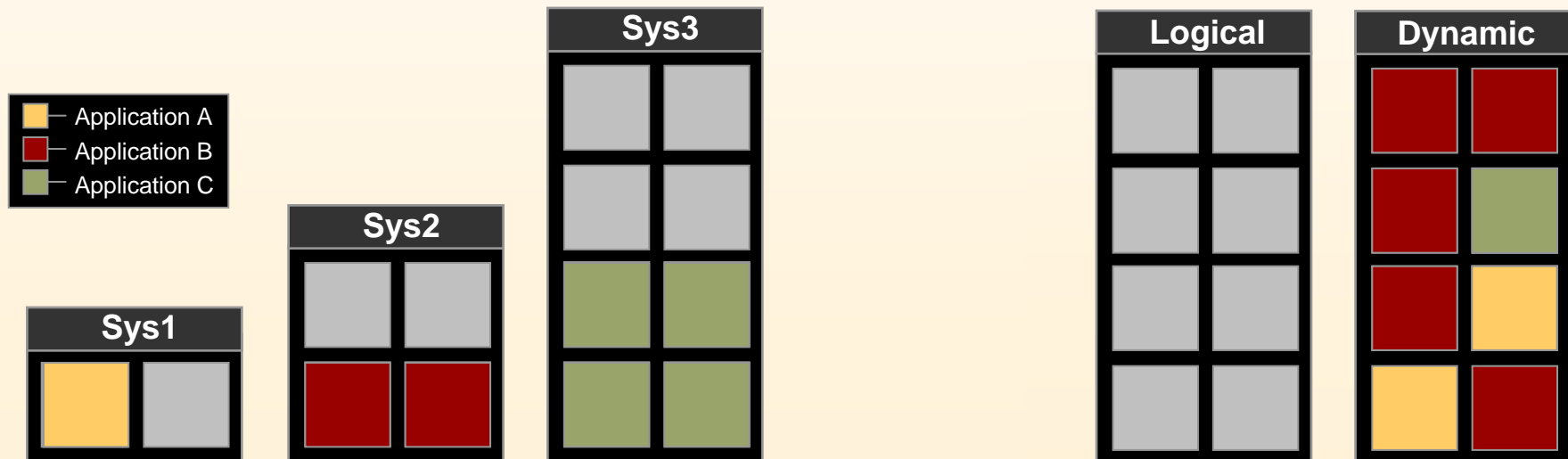


- Utilizes unused execution unit cycles
- Presents symmetric multiprocessing (SMP) programming model to software
- Natural fit with superscalar out-of-order execution core
- Dispatch two threads per processor: *“It’s like doubling the number of processors.”*
- Net result:
 - **Better performance**
 - **Better processor utilization**

IBM partitioning innovations help solve this problem

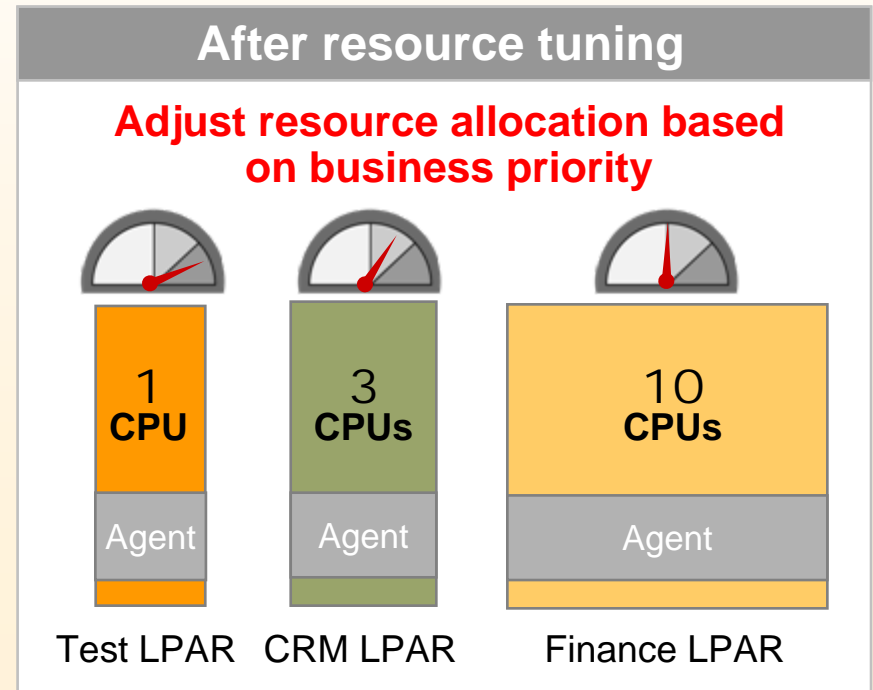
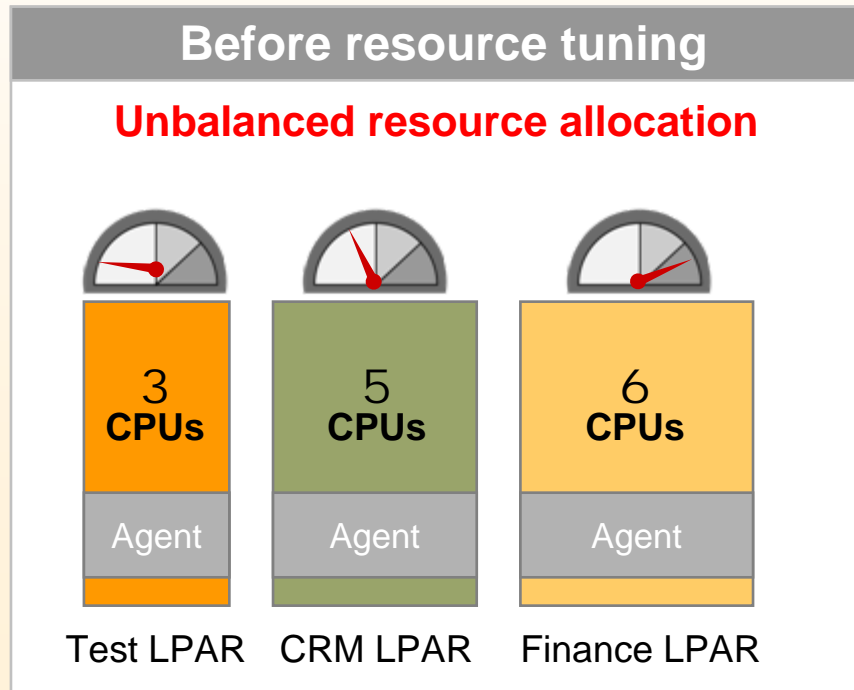
In 2001: Logical Partitioning **enabled consolidation of multiple application workloads**

In 2002: Dynamic Logical Partitioning, **enabled dynamic reassignment of workloads**



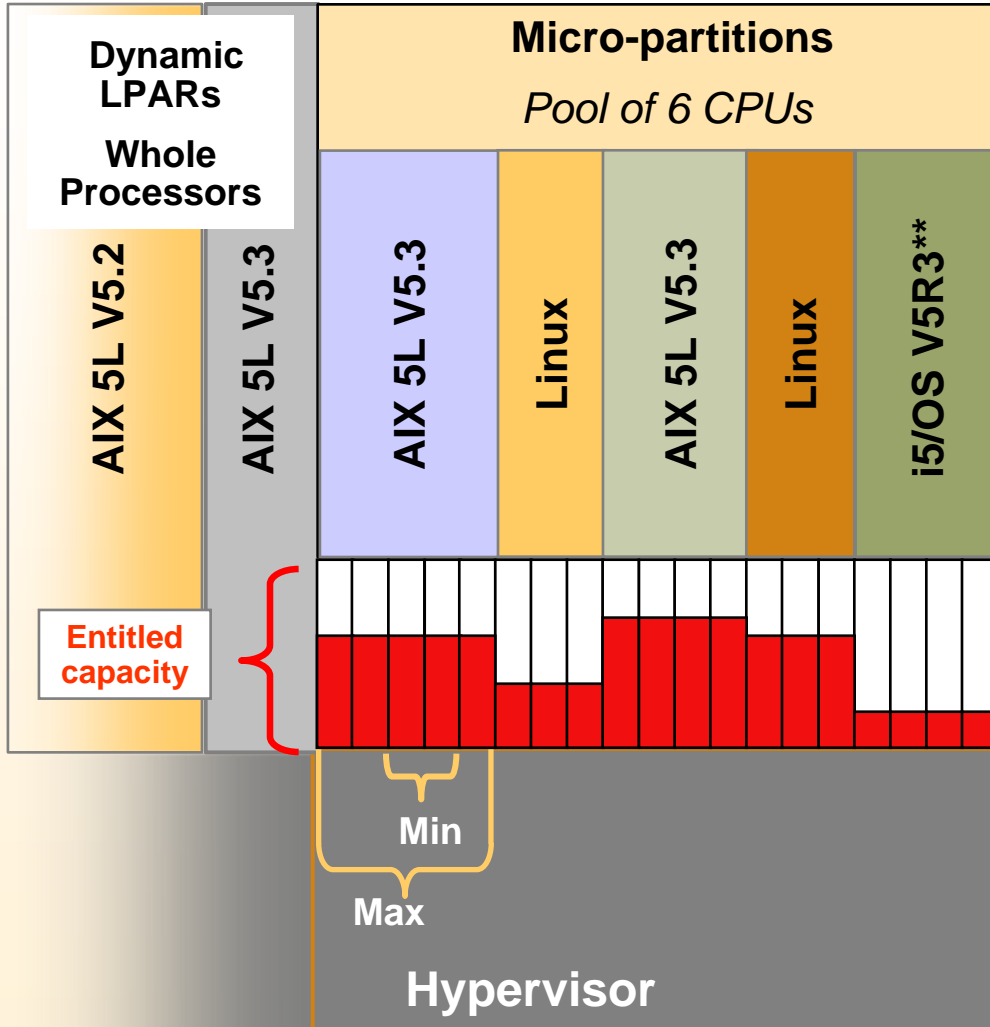
Meeting Demand: Utilizing Available CPU Cycles

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



Micro-Partitioning technology

AIX 5.3/AIX 6/Linux/i5 & iOS



Micro-Partitioning technology allows each processor to be subdivided into as many as 10 “virtual servers”, helping to consolidate UNIX® and Linux applications.

Partitioning options

- Micro-partitions: Up to 254*
- Dynamic LPARs: Up to 32*
- Combination of both

Configured via the HMC

Number of logical processors

- Minimum/maximum

Entitled capacity

- In units of 1/100 of a CPU
- Minimum 1/10 of a CPU

Variable weight

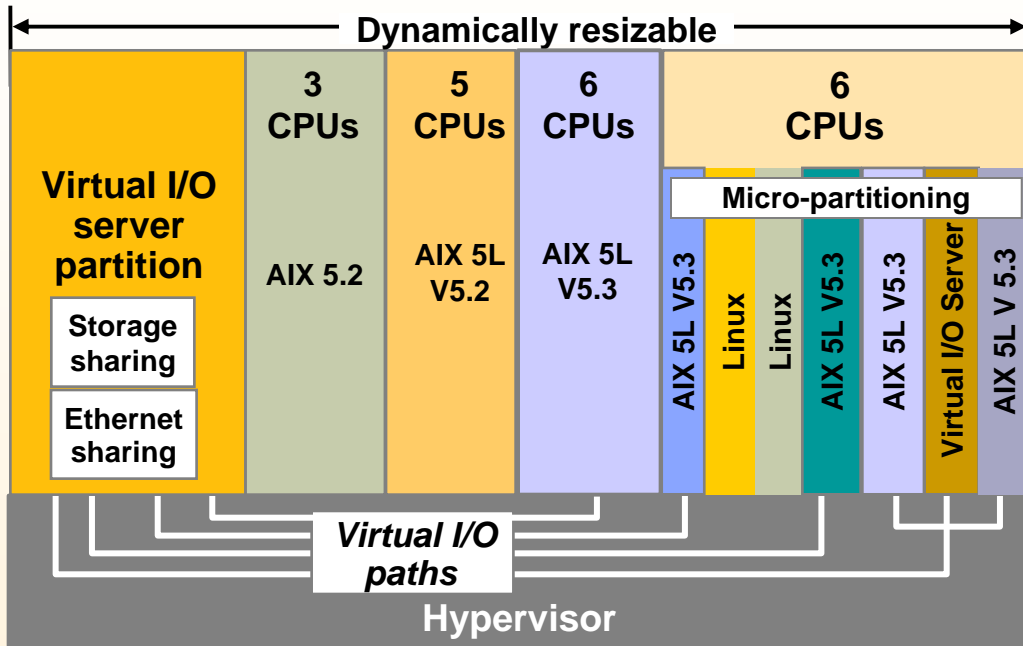
- % share (priority) of surplus capacity

Capped or uncapped partitions

*on p5-590 and p5-595

** on p5-570, p5-590, and p5-595

More Efficiency: CPU & I/O Sharing



Included features:

Enablement for Micro-Partitioning

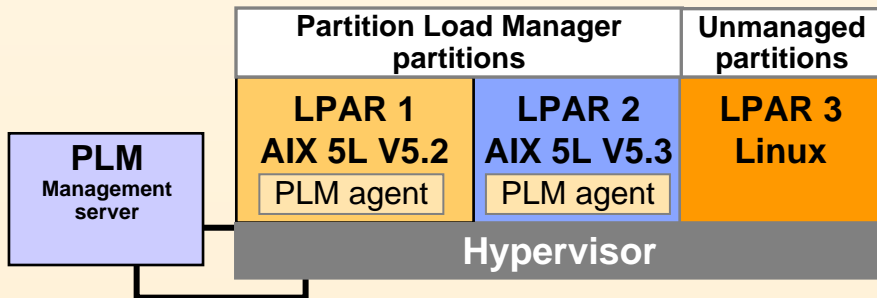
- Share processors across multiple partitions
- Minimum partition 1/10th processor
- AIX 5L V5.3 or Linux*

Virtual I/O server

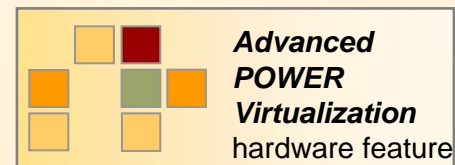
- Shared Ethernet
- Shared SCSI and Fibre Channel-attached disk subsystems
- Supports AIX 5L V5.3 and Linux*

Partition Load Manager

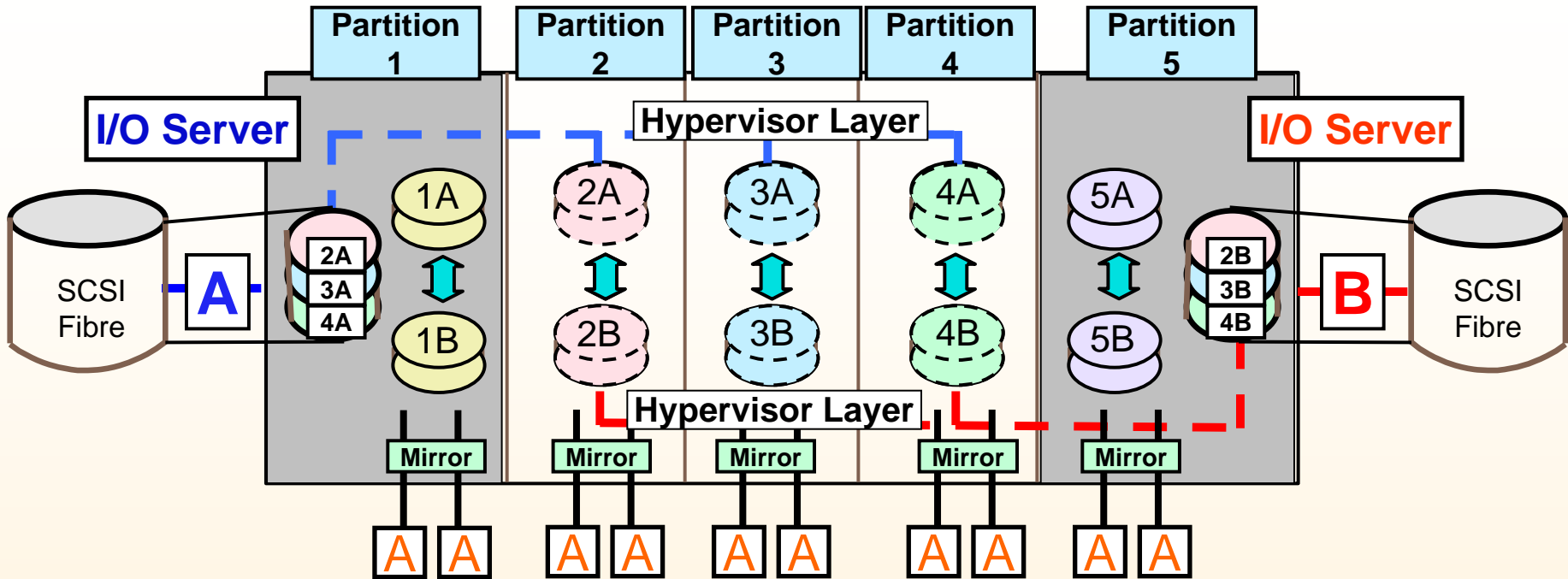
- Both AIX 5L V5.2 and AIX 5L V5.3 supported
- Balances processor and memory request



* SLES 9 or RHEL AS 3



POWER 5 & 6 / AIX 5.3 LPAR DISK Sharing



Using LVM mirroring for Availability

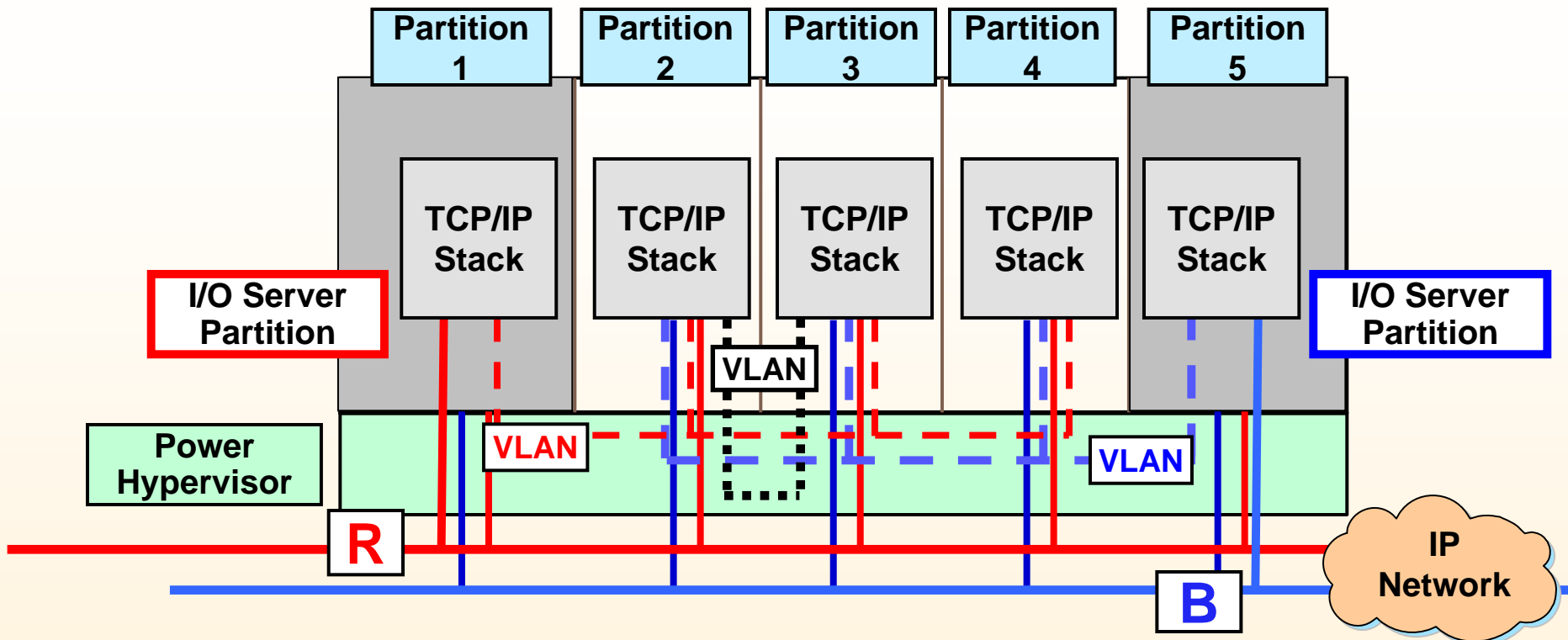
One physical drives appears to be multiple logical drives

- LUNs appears as individuals logical drives

Minimizes the numbers adapters

SCSI and Fibre supported

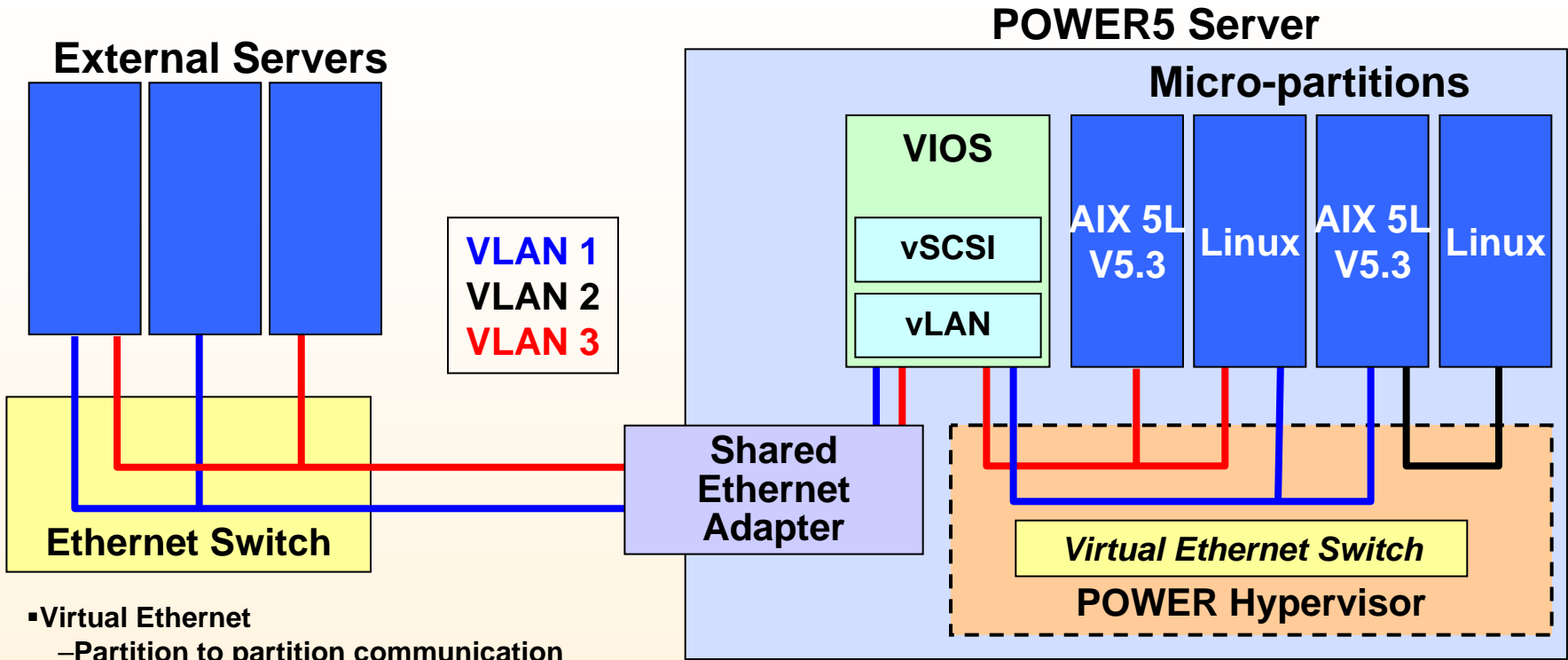
POWER5 / AIX 5.3 LPAR Virtual Ethernet



- Configured like a standard Ethernet
- IP forwarding provided by I/O Server partition
- Can have multiple connections per partition
- Inter-LPAR communications: IPv4 & IPv6 plus large MTUs

Virtual networking

Virtual Ethernet helps reduce hardware costs by sharing LAN adapters



- Virtual Ethernet
 - Partition to partition communication
 - Requires AIX 5L V5.3 and POWER5
- Shared Ethernet Adapter
 - Provides access to outside world
 - Uses Physical Adapter in the Virtual I/O Server
- VLAN – Virtual LAN
 - Provide ability for one adapter to be on multiple subnets
 - Provide isolation of communication to VLAN members
 - Allows a single adapter to support multiple subnets

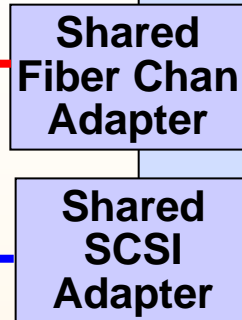
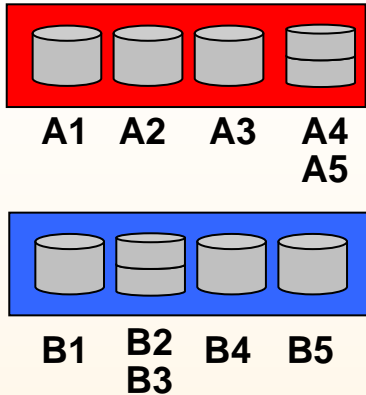
- IEEE VLANS
 - Up to 4096 VLANS
 - Up to 65533 vENET adapters
 - 21 VLANS per vENET adapter

Available via optional Advance POWER Virtualization or POWER Hypervisor and VIOS features.

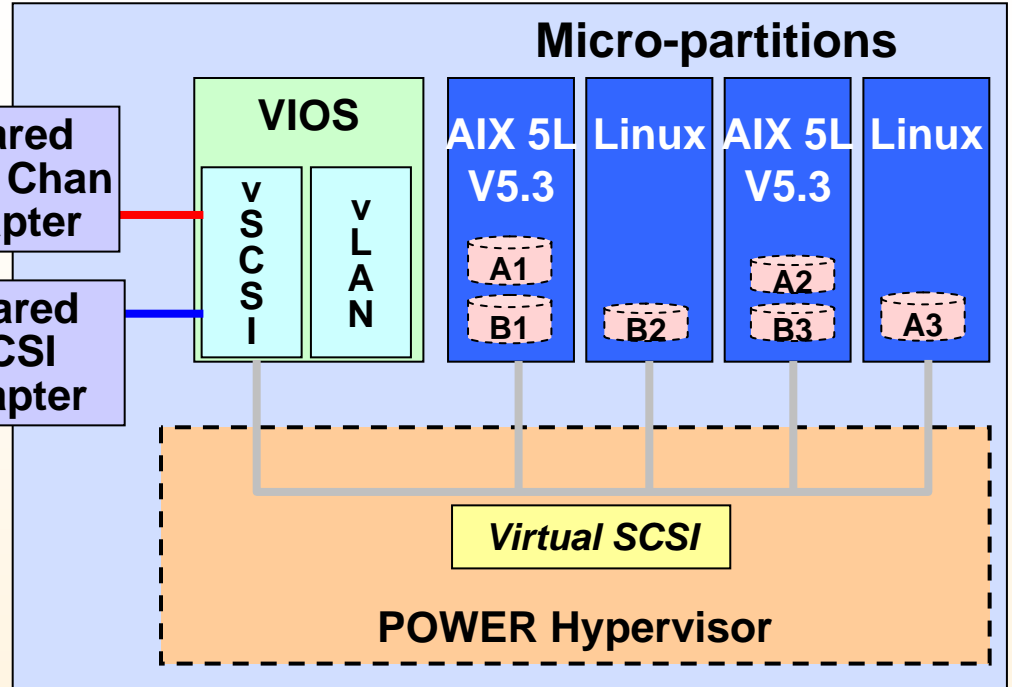
Virtual SCSI

Virtual I/O helps reduce hardware costs by sharing disk drives

External Storage



POWER5 Server



VIOS owns physical disk resources

- LVM based storage on VIO Server
- Physical Storage can be SCSI or FC
- Local or remote

Micro-partition sees disks as vSCSI (Virtual SCSI) devices

- Virtual SCSI devices added to partition via HMC
- LUNs on VIOS accessed as vSCSI disk
- VIOS must be active for client to boot

Multiple LPARs can use same or different physical disk

- Configure as logical volume on VIOS
- Appear a hdisk on the micro-partition
- Can assign entire hdisk to a single client

Available via optional Advance POWER Virtualization or POWER Hypervisor and VIOS features.

On Demand Capabilities: Keeps on Expanding

- **CoD Offerings (Capacity on Demand)**
 - Trial
 - Reserve
 - On/Off
 - Backup
- **RAS (Reliability Availability Serviceability)**
 - Processor and Memory Sparing
- **Building Block Architecture**
 - Pay as you grow

Capacity on Demand

Capacity Upgrade on Demand

- Upgrade system with processors and/or memory
- No special contracts, no required monitoring (no ability to turn off the capacity)
- Purchase agreement

On/Off Capacity on Demand

- Temporary use of requested number of processors or amount of memory
- Client selects the capacity and activates the resource (registered system)
- Capacity can be turned on and off by the client
- Information captured by IBM (or reported to IBM)
- Rental agreement

Reserve Capacity on Demand

- Processor resources only (processor days)
- Capacity can be turned on and off by the client
- Prepaid debit agreement
- Requires AIX 5L V5.3 and APV

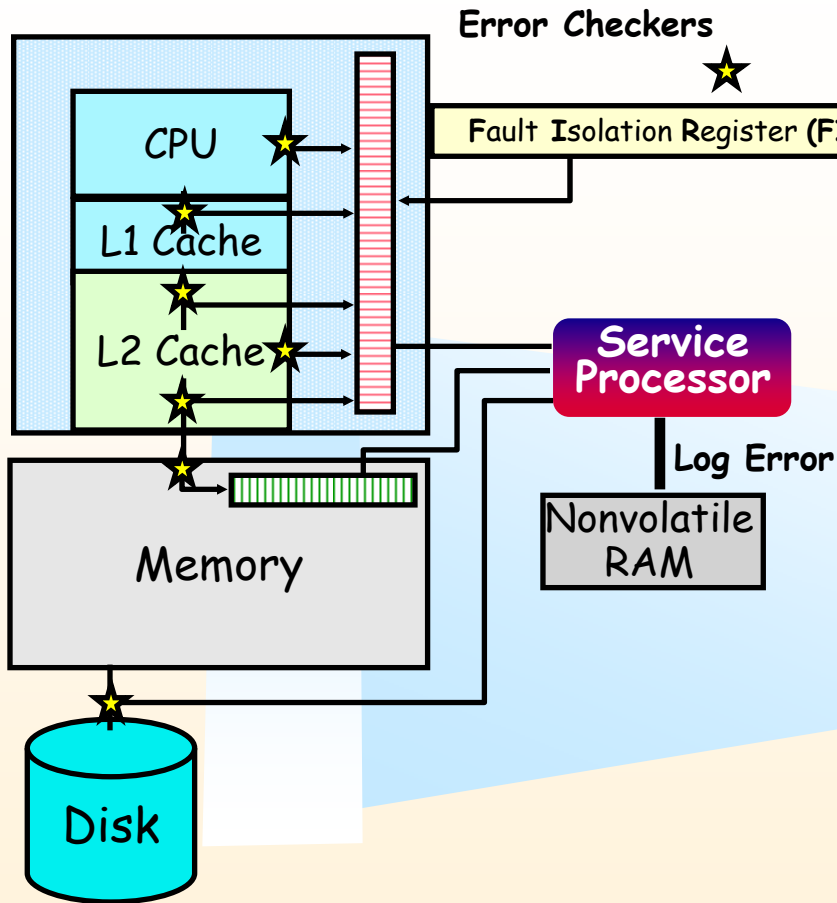
Trial Capacity on Demand

- Allow clients to test the effects of additional processors and/or memory
- Partial or total activation of processors and memory
- Resources available for fixed time
- No formal commitment required

Dynamic Processor Sparing

- Automated replacement of de-allocated processors
- Unassigned or inactive processors

First Failure Data Capture



Competitive Strategy - Failure Re-create

- f* Run diagnostic testcases during service call
- f* Repair based on testcase symptom
- f* Questionable correlation to original customer problem
- f* Development focus on improved testcases
- f* Open service action plan if failure not re-created
- f* Testcases used to isolate failures and verify correct operation

First Failure Data Capture

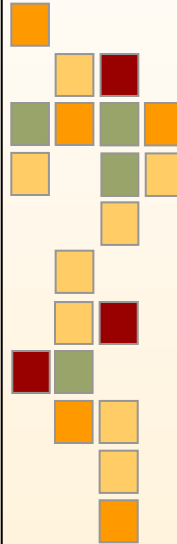
- f* Specialized hardware designed to capture failure data at the time of failure
- f* Repair based on root-cause analysis
- f* Direct correlation to original problem
- f* Engineering focus on built-in error detection and capture
- f* Service action plan driven by captured failure information
- f* Testcases used only to verify operation

Reliably identify failing component
reducing costly downtime

POWER RAS additions...

Primary POWER RAS features

- ✓ First Failure Data Capture
- ✓ DDR Chipkill™ memory
- ✓ Bit-steering/redundant memory
- ✓ Memory soft scrubbing
- ✓ Redundant power, fans
- ✓ Dynamic Processor Deallocation
- ✓ Dynamic processor sparing
- ✓ ECC memory
- ✓ Persistent memory deallocation
- ✓ Hot-plug PCI slots, fans, power
- ✓ Internal light path diagnostics
- ✓ Hot-swappable disk bays
- ✓ I/O error handling extended beyond base PCI adapter
- ✓ ECC extended to inter-chip connections for the fabric/processor buses



2005 RAS Enhancements

- **Service Processor Failover:**
 - p5-570, p5-590 and p5-595
 - New service processor option for p5-570
 - ✓ Required for “Failover” support
- **Dynamic Firmware Maintenance:**
 - Apply firmware fixes without system disruption
 - Fixes only, not new functionality
- **Hot I/O Drawer Add**
 - Install remote I/O drawers without system disruption

HMC required to enable these functions..

Enable high system availability – Building on world-class hardware RAS

Summary of key pSeries RAS features

Core System Design

- High quality parts
- Fewer parts = Fewer failures
- Designed for low power consumption (less heat = fewer failures)
- Manufacturing methods, packaging, cooling
- Continuous System and Commodity Quality Actions
- Integrated RAS features
- Failure Avoidance Methodology
- Designed for Ease of Service

Fault Resilience

- N+1 Power Supplies, regulators, power cords
- Dual redundant fans
- Dynamic Processor Deallocation and sparing
- "Chipkill" Technology
- Predictive Failure Analysis
- Auto Path Reassignment - data paths, power

System Restore

- Deferred Repair
- Concurrent Repair
- LED Service Identification
- Service Consoles
- Migration to Guided Maintenance

Fault Isolation & Diagnosis

- First Failure Data Capture
- Run Time Self Diagnostics
- Service Processor
- Rifle-shot repairs (no "plug and pray" parts replacement approach)

Summary

Why IBM System p AIX/LINUX solutions

- IBM has a **Proven** Technology
- IBM has a **Consistent** Product line
- IBM has a **Clear Direction**
 - A good hockey player plays where the puck is. A great hockey player plays where the puck is going to be.
 - Wayne Gretzky
- IBM wants to **EARN** your business