IBM Systems and Technology Group University 2005

IBM Systems and Technology Group University 2005

v 1.0

IBM Systems and Technology Group University 2005

Performance Considerations With Micro Partitioning

Course #: PT06

Harold Lee
Advanced Technical Support - Americas

v 1.0

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, DBE, e-business logo, ESCO, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/30, VM/ESA, VSE/ESA, Websphere, xSeries, z/OS, zSeries, z/VM

The following are trademarks or registered trademarks of other companies

Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation
Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., 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 avalable in your area.

All statements regarding IBM's future direction and intent are subject to change or withdraw al 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.

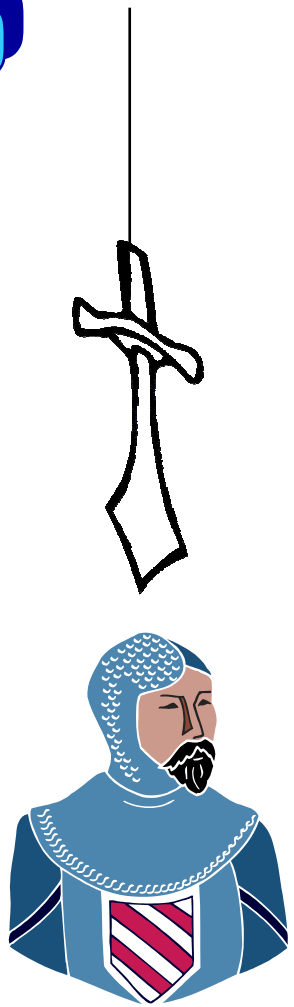
Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

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.

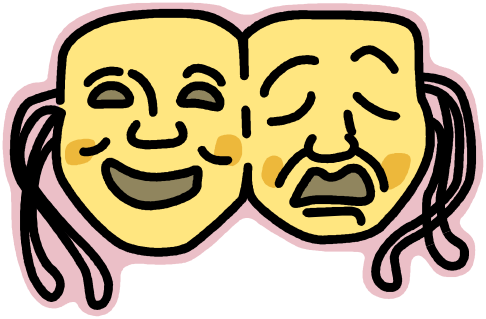
WHY Micro-partition?

- **Server Consolidation**
  - Many to one Micro-partition
  - Consideration – Software Licensing
- **Server Provisioning**
  - Rapid deployment for unplanned demands
  - Creating test environment
- **Virtual Server Farms**
  - Shared resources
  - Scalability



WHY Micro-partition?

The virtualization of processors is employed to promote flexibility in using the system and employing fractional processing power but incurs some overhead.



Setting Expectations

- **Addressed - Performance CONSIDERATIONS with Micro-Partitioning**
  - How subsystems interact
  - How virtual components behave
  - How some virtual components perform
- **Not Addressed – Micro-Partitioning has XX% of overhead**
  - Extremely OS, application, and load dependant
  - Empirically determined (benchmark)…. Perhaps

- History of Technology
- Overview of Virtual Components
- Performance Considerations of Components
- Performance Tools
- Summary
- Additional Information & Resources

History of Technology

- **1974 - Dr. John Cocke – RISC Technology**
- **1978-83 – 801 processor – 1 instruction/clock**
  - Pipelining
  - IBM Office Products in Rochester – Processor
  - Research OPD Micro Processor - ROMP
- **1986 - Technology Transfer – Austin - RTPC**
  - 2.3 cycles/instruction
  - 4.3 MIPS

History of Technology

- **1991 – RISC System/6000 – Superscalar workstation**
  - 3 instructions/cycle
    - Instruction Stream Processor
    - Fixed-Point Processor
    - Floating-Point Processor
  - Complex Optimizing Compilers
- **1993 – Wildhorse SP1 Scalable/Parallel**
- **1994-1995 – AIX 4 – SMP**
  - Threads VS Process
  - SMP
  - Divergence of PWR2 technology

History of Technology

- **1997 - 1998 – AIX 4.3 – S70**
  - AIX 4.3
    - 64 bit technology
    - Posix threads
  - Raven – 64 bit Power PC – Apache Chip
  - Blackbird – 64 bit power PC – Northstar Chip
- **1999 – AIX 4.3.3**
  - Workload Manager

History of Technology

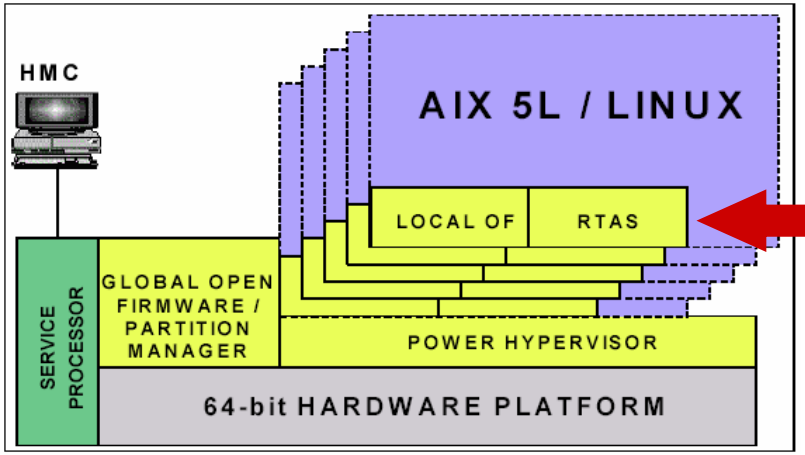
- **2001 – Power4 “Gigaprocessor”, AIX 5L V 5.1**
  - Multiple processors on single chip
  - Regatta 8-32 way SMP
  - Increased WLM features
  - Hardware multi-threading
  - Hypervisor, LPAR Technology introduced
- **2002 – 2003 Power4+, AIX 5L V 5.2**
  - Dynamic LPAR
  - CUOD

History of Technology

- **2004 – Power5, AIX 5.3**
  - Micro-Partitioning
  - Intra-partition
  - Virtual I/O
  - Virtual Ethernet
  - Simultaneous Multi-threading

Overview of Virtual Components

- **Vitrualization Technology – The pooling of system resources via the POWER Hypervisor to access processors, memory, and I/O devices across logical partitions**
- **POWER Hypervisor – The underlying control mechanism that resides below the operating systems and above the hardware. It owns all of the of the resources and creates partitions by allocating resources and sharing them**



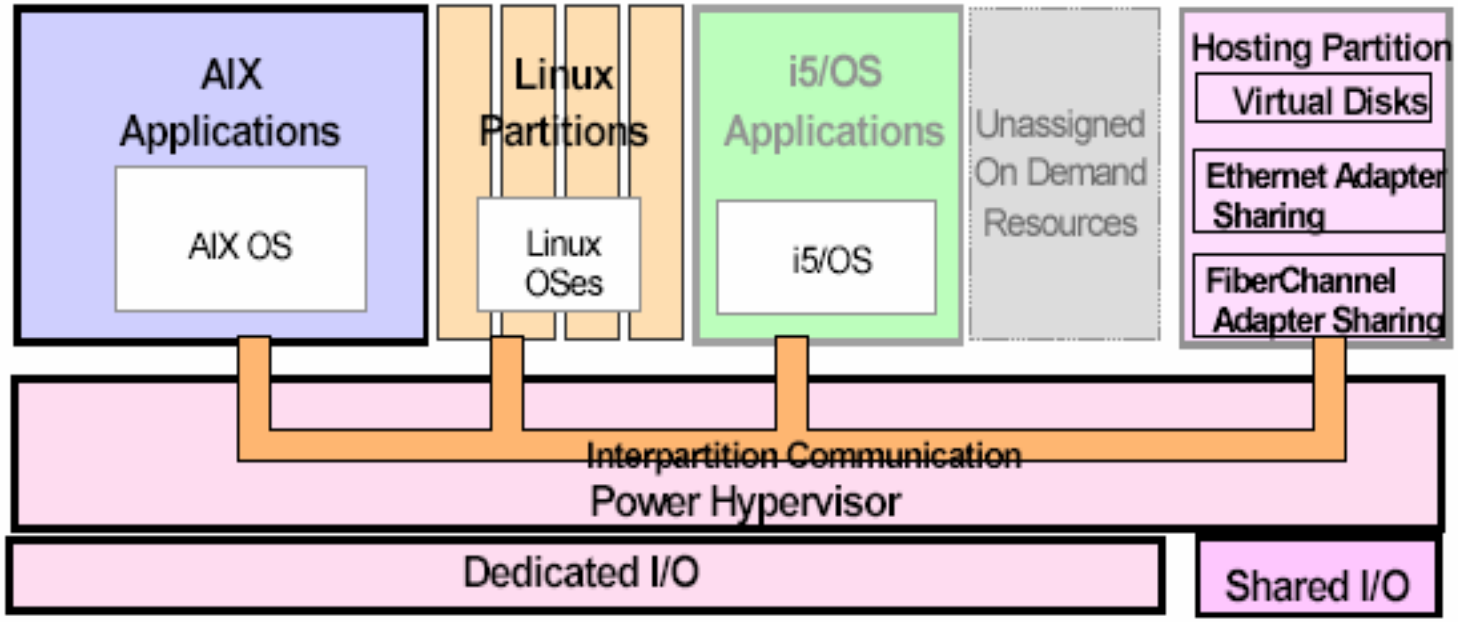
Real-time Abstraction Services

Overview of Virtual Components

- **Virtual I/O – 5 different concepts**
  - Virtual SCSI
  - Virtual Ethernet
  - Virtual Serial
  - Virtual I/O Server – Special LPAR
  - Virtual network to real network mechanism – Shared ethernet adapter

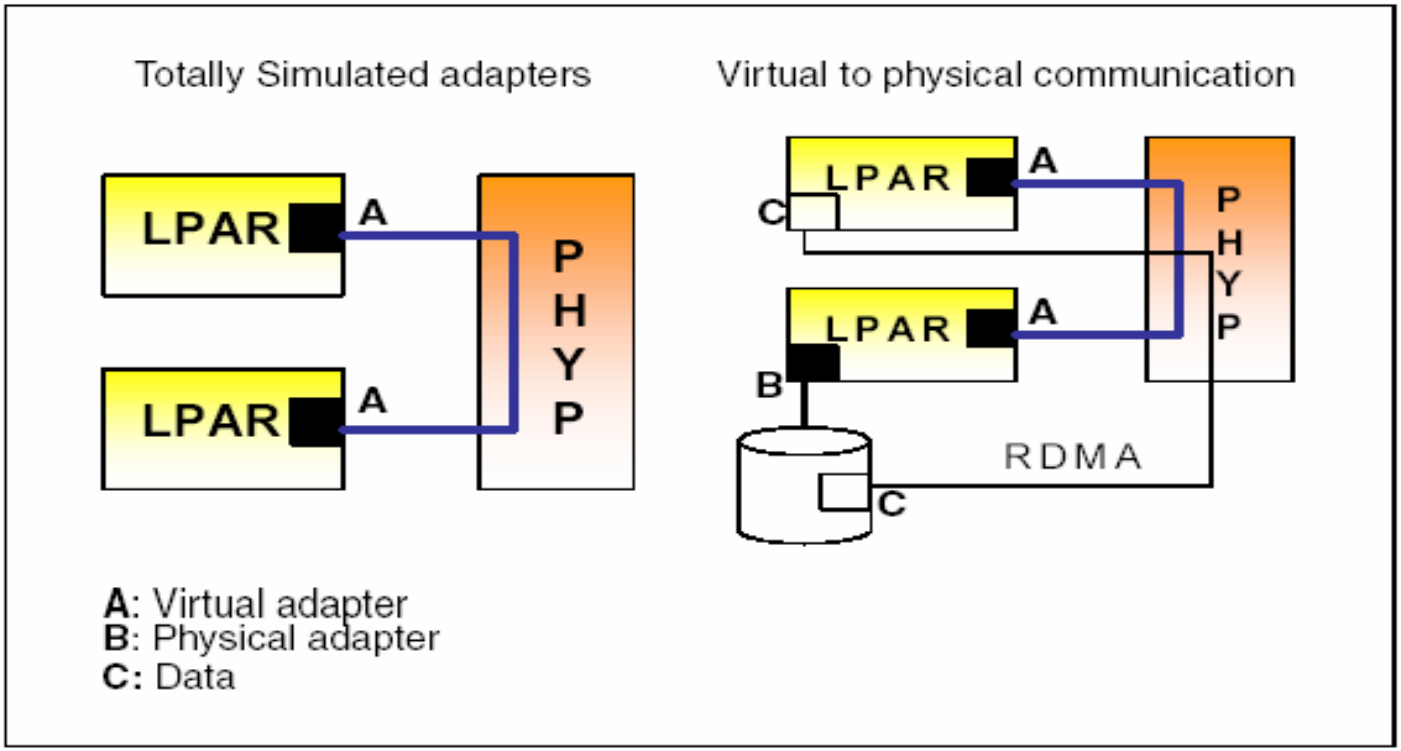
Overview of Virtual Components

- **Virtual I/O**



Overview of Virtual Components

- **Virtual I/O**

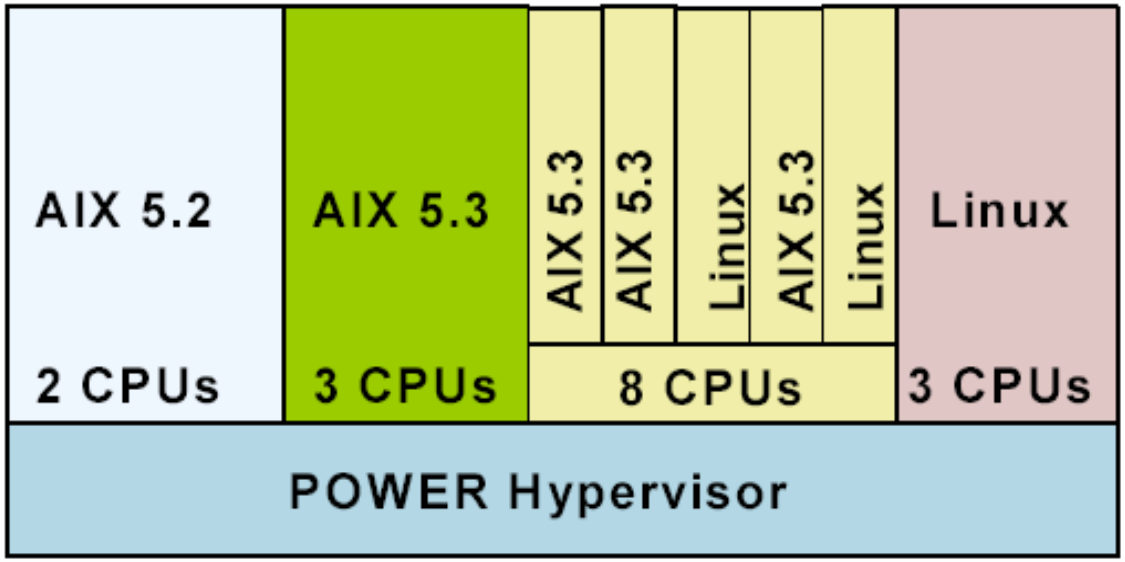


Overview of Virtual Components

- **Dedicated Processor Partitoning**
  - Processor is owned by a partition and is not shared
  - LPAR capacity is limited to dedicated processor(s)
  - LPAR can use combination of physical and virtual I/O resources
  - LPAR with dedicated processors can still use virtual I/O that is shared by other partitions including virtual partitions.

Overview of Virtual Components

- **Micro-Partitioning**
  - Physical processors abstracted into virtual processors
  - Vary from .1 processor in .01 increments

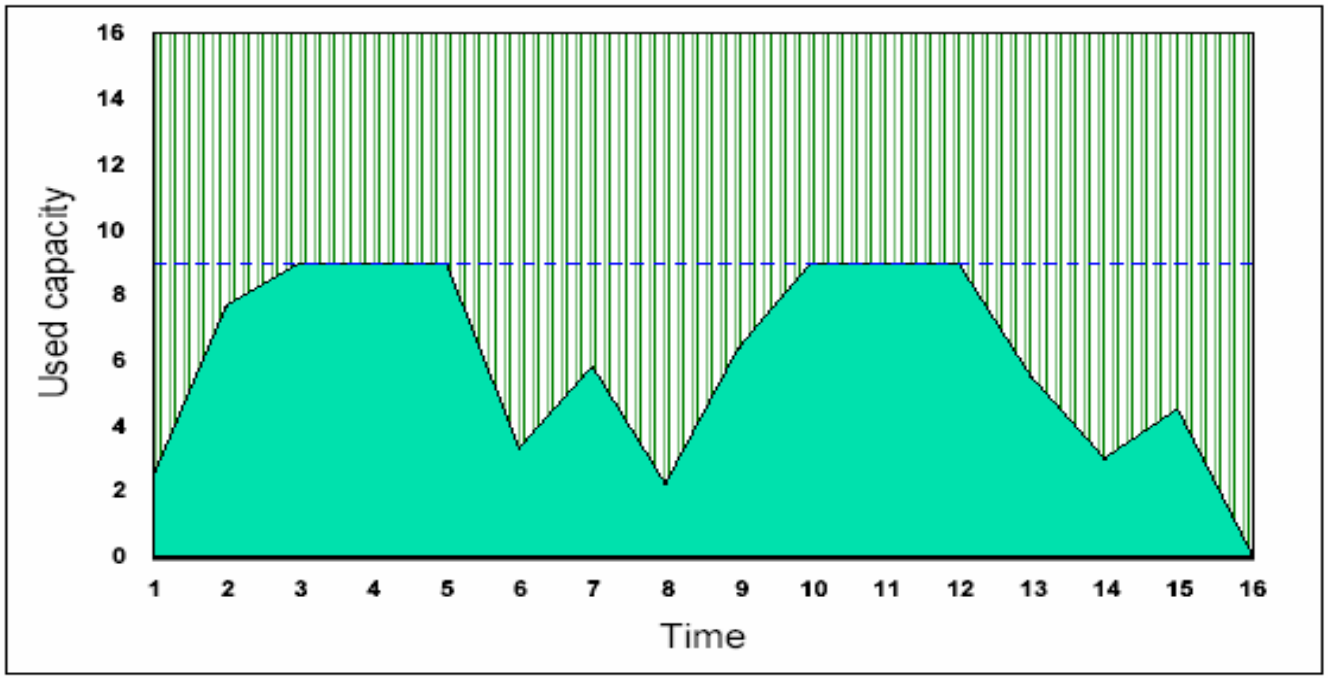


Overview of Virtual Components

- **Micro-Partitioning**
  - New Terminology
    - Processing Unit
    - Layers Of Processor Abstraction
      - Physical Processor
      - Virtual Processor
      - Logical Processor
    - Processor Pools
  - Types of Micro-partitions
    - Capped – Hard maximum limit even with excess capacity
    - Uncapped – Soft maximum limit. May get more than entitlement

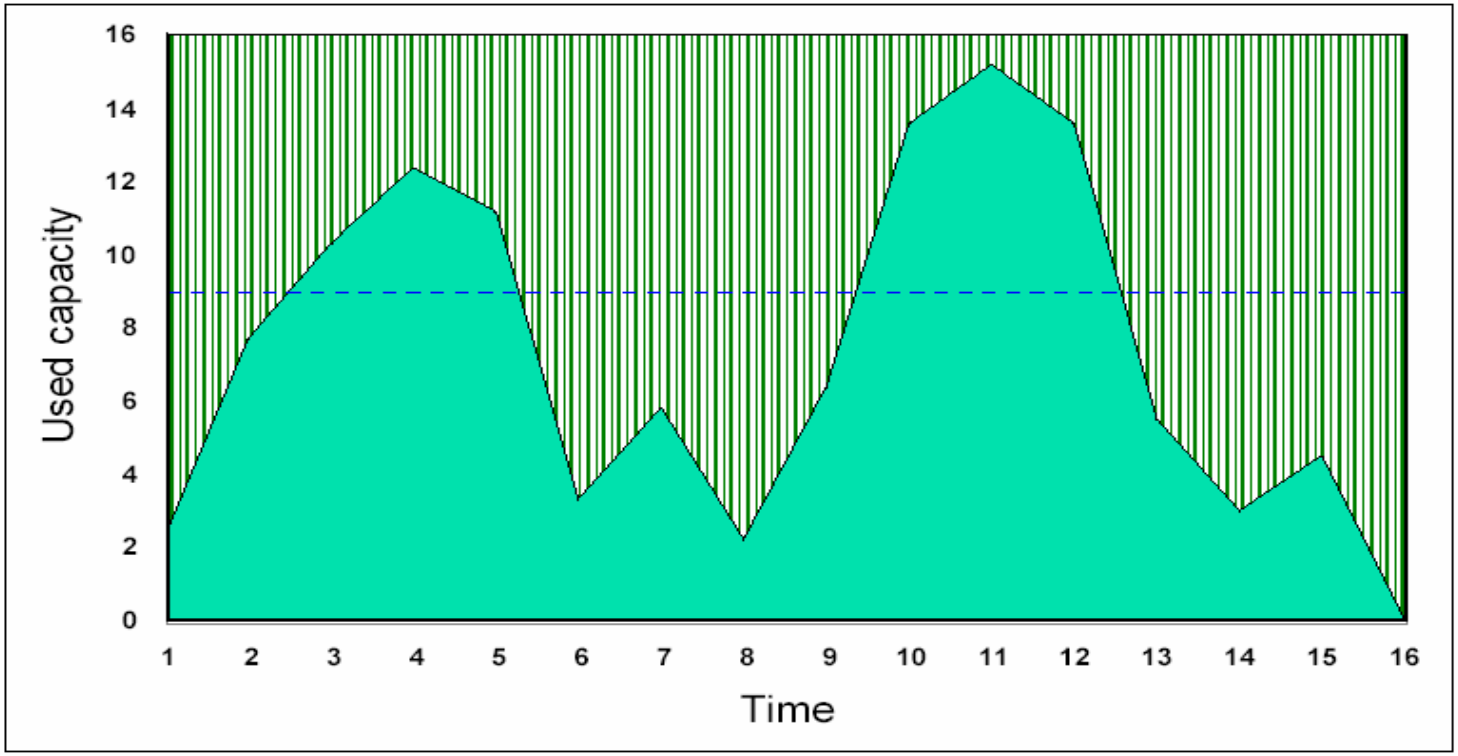
Overview of Virtual Components

- **Micro-Partitioning - Capped**



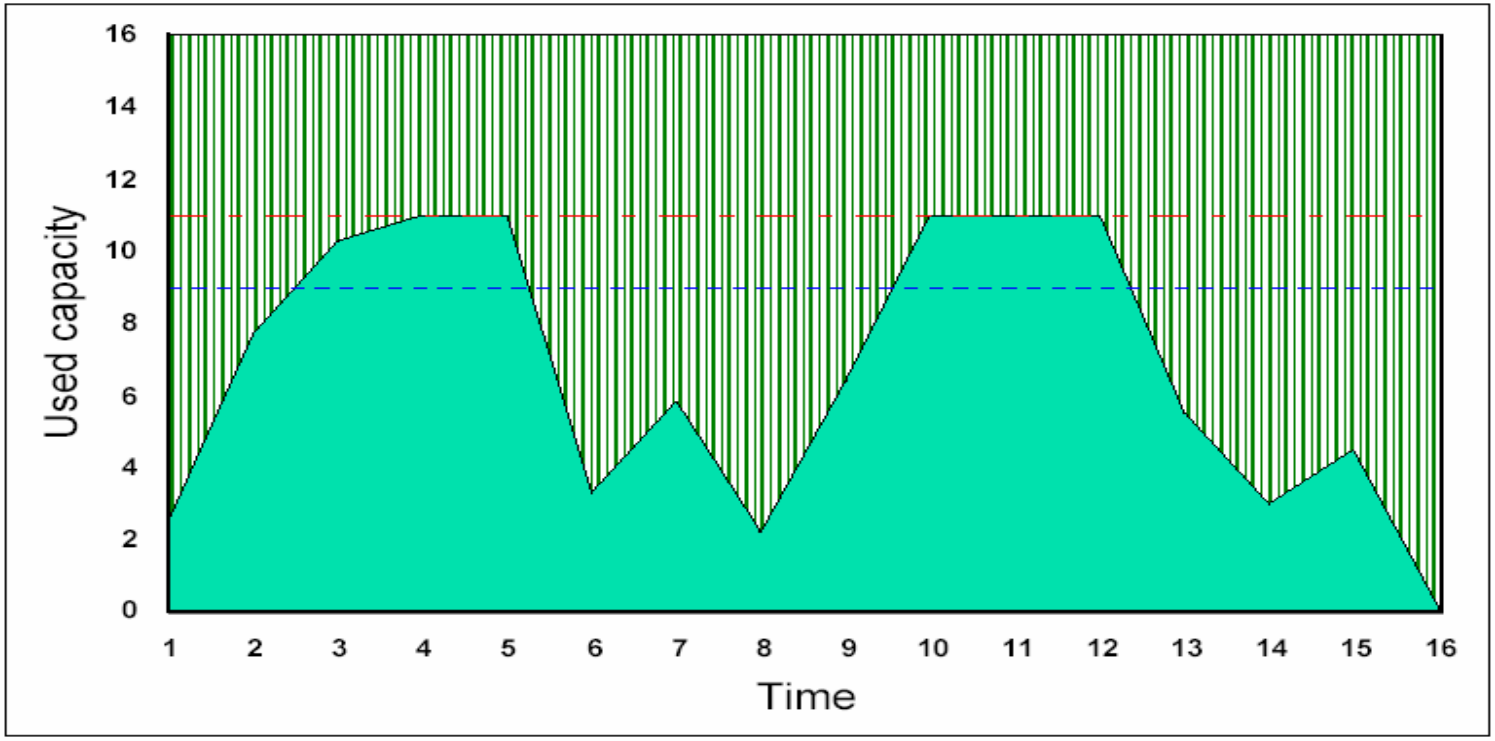
Overview of Virtual Components

- **Micro-Partitioning - Uncapped**



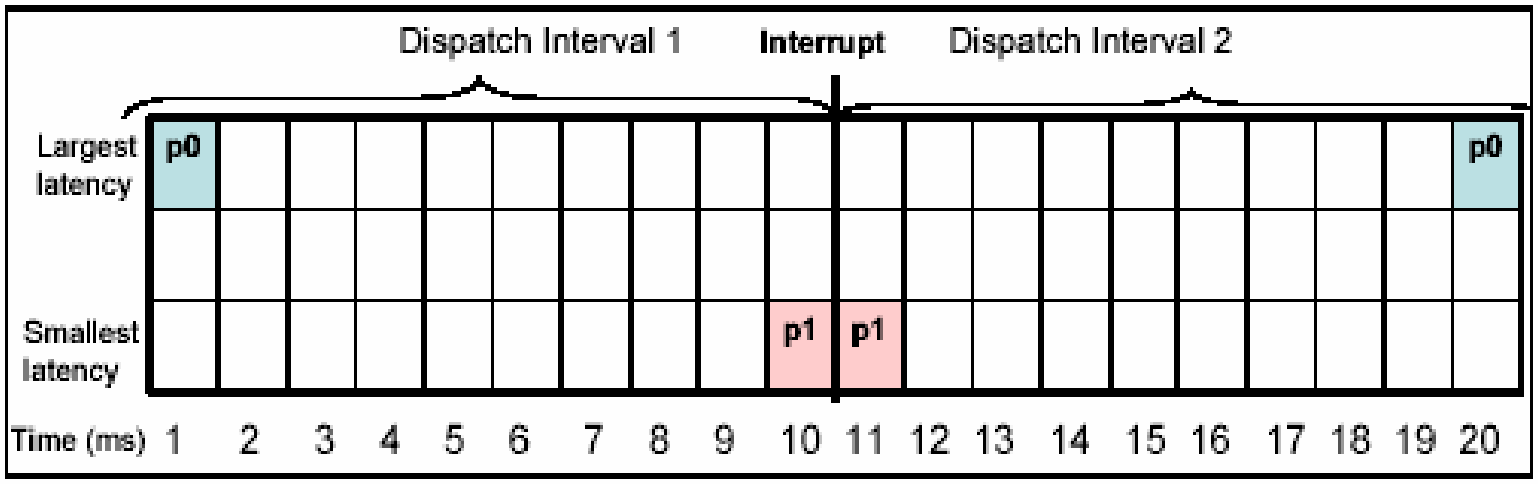
Overview of Virtual Components

- **Micro-Partitioning - Uncapped weighting**
  - Variable capacity weight – value between 0 and 255 indicating relative share weight



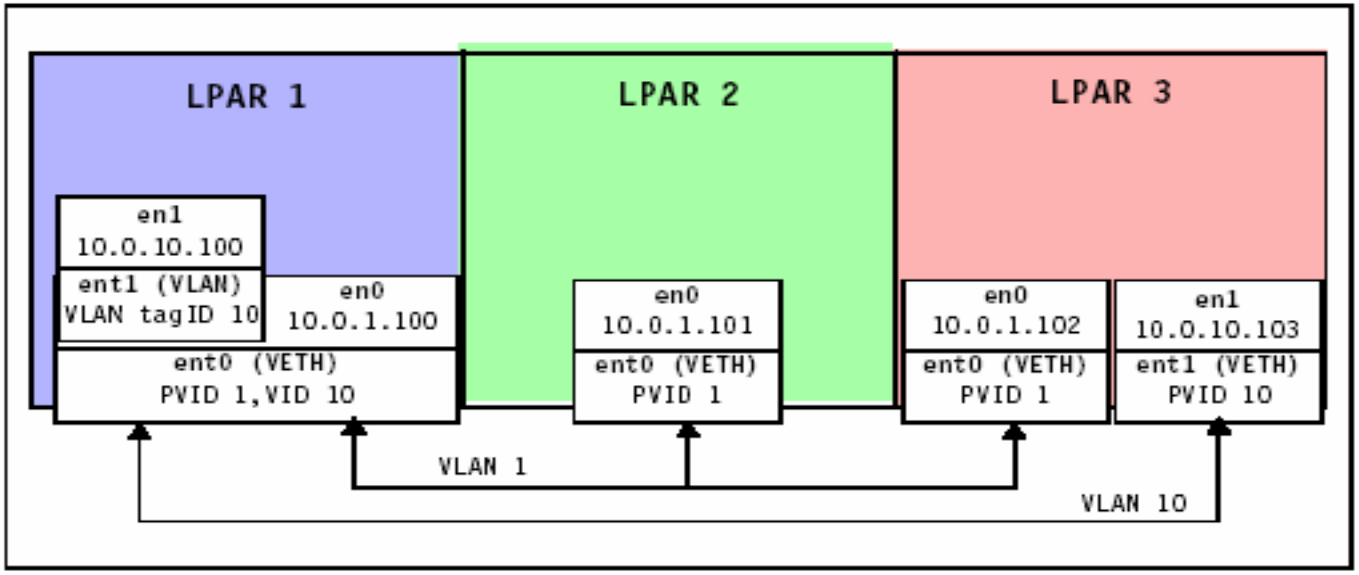
Overview of Virtual Components

- **Hypervisor Dispatch of Micro-partitions**
  - Hypervisor will attempt to dispatch partitions every 10 milliseconds
  - Dispatch is not guaranteed to be sequential
  - Dispatch latency can be from 0 to 18 Milliseconds



Overview of Virtual Components

- **Virtual Ethernet – Virtual LANS – VLANS**
  - Do not confuse with network-topology VLANS
  - Logical connection between one or more partitions
  - No physical adapters required – Memory based
  - May be bridged to external LAN through Virtual I/O Server LPAR



Overview of Virtual Components

- **Virtual Ethernet – Virtual LANS – VLANS**
  - Throughput is a function of Micro-partition entitlement and MTU size
  - Relative throughput scales linear with entitlement
  - Relative throughput increases with MTU size
  - CPU consumption increases with MTU size
  - Overall performance is maximized with SMT turned on
  - If VLAN communicates with Virtual I/O server set MTU sizes to match
  - Keep tcp_pmtu_discover to default (active discovery)

Overview of Virtual Components

- **Virtual Ethernet – Virtual LANS – VLANS**

Overview of Virtual Components

- **Virtual Serial**
  - May only be used for providing virtual console to partitions
  - Cannot be used for intra partition communications
  - Cannot be used for any other purpose (HACMP heartbeat)
  - No specific performance considerations.

Performance Monitoring Commands

COMMAND	lparstat	mpstat	vmstat	iostat	sar	topas	xmperf
Measurement							
CPU	YES	YES	YES		YES	YES	YES
Memory			YES			YES	YES
Disk I/O			YES	YES	YES	YES	YES
Network					YES		YES
Hypervisor	YES						
Environment							
Dedicated Partition	YES	YES	YES	YES	YES	YES	YES
Micro-partition	YES	YES	YES	YES	YES	YES	YES
SMT	YES	YES	YES	YES	YES	YES	YES
I/O Server			YES	YES	YES		

Performance Monitoring Commands

Command	Function	Main measurement
lparstat Modified	Logical partition information and statistics	CPU, Hypervisor
mpstat NEW	Physical and logical processors statistics	CPU
vmstat Modified	CPU and virtual memory monitoring	CPU, memory
iostat Modified	System input/output device monitoring	Disk I/O
sar Modified	Physical, logical processors and I/O monitoring	CPU
topas Modified	Displays dynamically system statistics.	CPU, memory, I/O
xmperf Modified	Displays a great amount of system statistics	CPU, memory, I/O

- **Applications with response time criteria may not be good candidate for Micro-partitioning**
- **Applications without strong quality of service requirements are not good candidates for Micro-partitioning**
- **Applications that rely on polling may not be good candidates**
- **Applications with low average utilization with high peaks are good candidates for Micro-partitioning**
  - Mail servers
  - ERP
  - Web servers
  - Directory servers

- **Applications with high CPU usage and relatively constant demands may not be a good candidate for Micro-partitioning**
  - High performance computing – HPC
  - Decision support systems – DSS
  - Consider dedicated processor LPAR

Additional Resources & References

- **IBM Redbook: Advanced Power Virtualization on IBM P5 Servers – Architecture and Performance Considerations – SG24-5768-00 – working draft**
- **IBM Redbook: Introduction to Advanced POWER Virtualization on IBM p5 Servers, Introduction and basic configuration, SG24-7940**
- **White Papers**
  - IBM p5 570 Server Consolidation Using POWER5 Virtualization
  - IBM p5 570 Workload Balancing Using POWER5 Virtualization
- **History - "27 Years of IBM RISC" http://www.rootvg.net/column_risc.htm**

- *Using the Virtual I/O Server*

  http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/info/iphb1/iphb1.pdf

- *Virtual I/O Server and Partition Load Manager Commands Reference*

  http://publib.boulder.ibm.com/infocenter/eserver/v1r2s/en_US/info/iphb1/commands/commands.pdf