

# Virtualization Capacity Management and Hollywood

Charles Johnson  
Metron-Athene, Inc.

August 16, 2013  
Session Number (14079)



# The Rundown



- Introduction
- Hollywood
- Virtualization
- Virtualization Capacity Management Challenges
- Virtualization Key Metrics
- The Script - Telling the story
- The End

# Introduction



- Principal Consultant with Metron-Athene
- Involved in the IT industry for over 30 years
- Involved in Performance and Capacity Management during complete career
- VMware and ITIL certified

# The Rundown

- Introduction
- **Hollywood**
- Virtualization
- Virtualization Capacity Management Challenges
- Virtualization Key Metrics
- The Script - Telling the story
- The End

# Hollywood Movies

- What makes a successful Hollywood movie?
  - Makes you ask questions
  - Keeps you engaged and holds your attention
  - A story is told and the viewer can relate to it in some way
  - The ending relates to the beginning

# Top 5 Grossing Movies of all time



- Gone with the Wind
- Star Wars
- Sound of Music
- E.T.
- Titanic

# Capacity Management for Virtualization Hollywood Movie Crew

- Director – Capacity Manager
- Cast
  - Servers (Including Mainframe)
  - VM's
  - Business Metrics
- Key Grip – Reports
- Script - Presentation



# The Rundown



- Introduction
- Hollywood
- **Virtualization**
- Virtualization Capacity Management Challenges
- Virtualization Key Metrics
- The Script - Telling the story
- The End



# Virtualization Platforms

- “x86”
  - VMware vSphere
  - Windows Hyper-V
  - Linux Xen
- Unix
  - AIX
  - Sun Solaris
  - HP
- Mainframe
  - z/VM

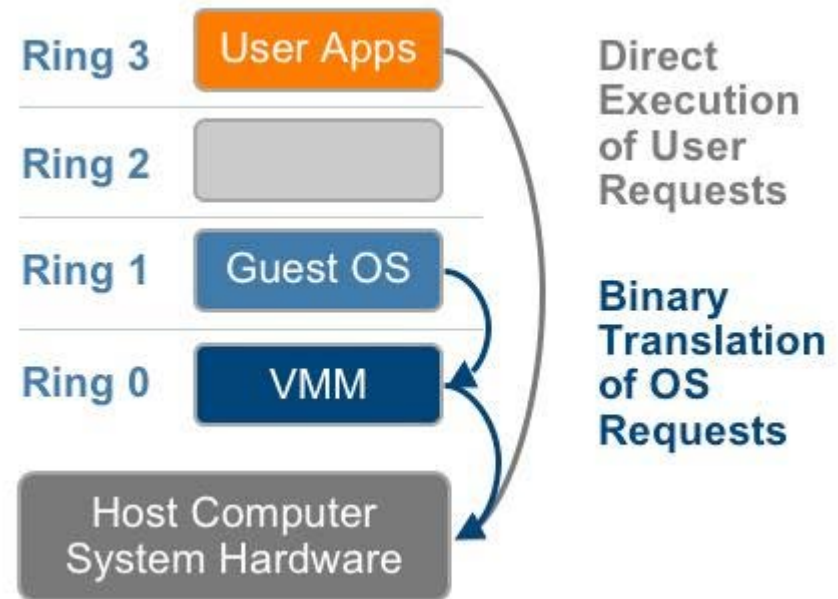
# Virtualization Overview



- Virtualization Techniques
  - Software Virtualization
  - Hardware Virtualization
- Memory Management
- Network Management
- Storage Management

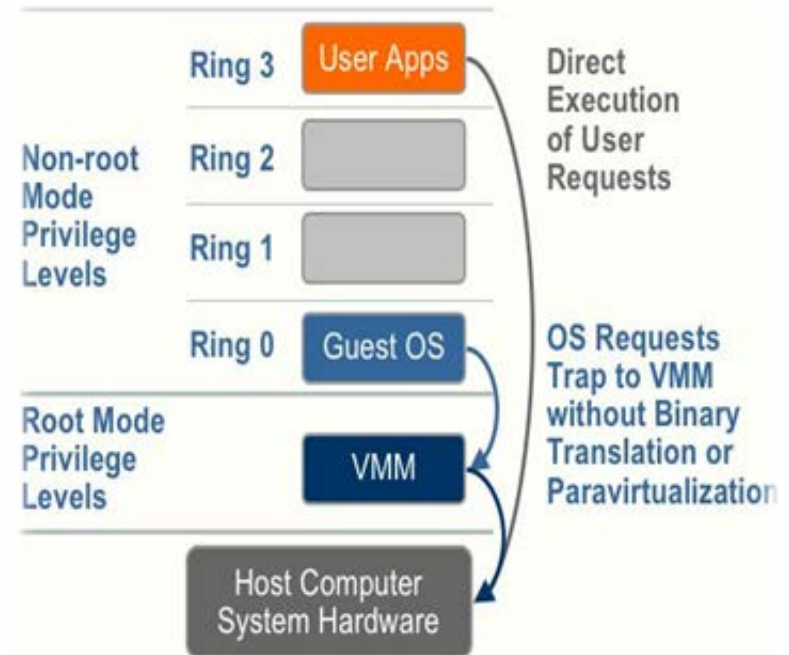
# Software Virtualization

- Original approach to virtualizing the (32-bit) x86 instruction set
- Binary Translation allows the VMM to run in Ring 0
- Guest operating system moved to Ring 1
- Applications still run in Ring 3



# Hardware Virtualization

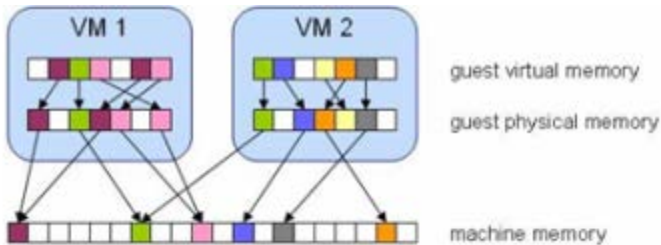
- Aim: is to simplify virtualization techniques
  - VMM removes Binary Translation whilst fully controlling VM.
  - Restricts privileged instructions the VM can execute without assistance from VMM.
- CPU execution mode feature allows:
  - The VMM to run in a root mode below 0
  - Automatically traps privileged and sensitive call to the hypervisor
  - Stores the guest operating system state in VM control structures (Intel) or blocks (AMD)



# Memory Management

- Memory management is in some cases the most critical component
- Processes see virtual memory
- Guest operating systems use page tables to map virtual memory addresses to physical memory addresses
- The MMU translates virtual addresses to physical addresses and the TLB cache help the MMU speed up these translations.
- Page table is consulted if a TLB hit is not achievable.
- The TLB is updated with virtual/physical address map, when page table walk is completed.

# vSphere Memory Management Features



Transparent page sharing

Memory borrowing

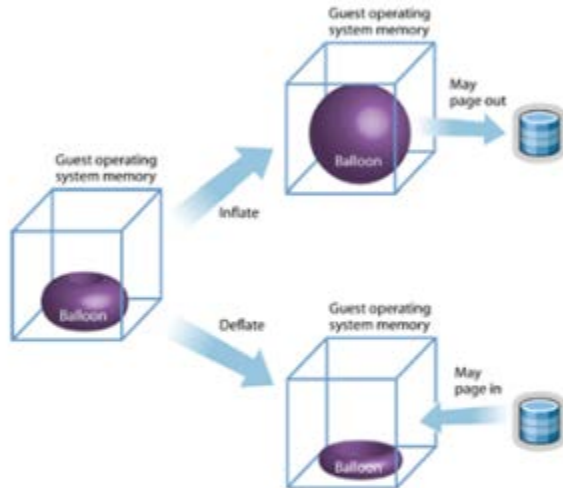
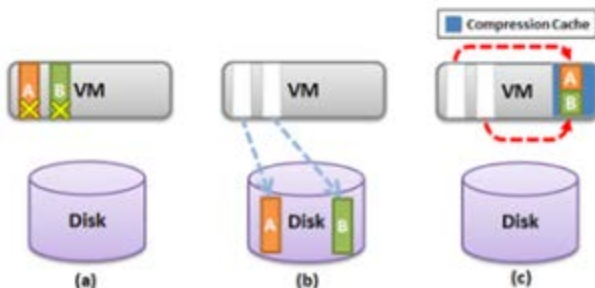


Figure 1: Memory balloon driver in action



Memory compression

...and Paging

# Network Management



- Reducing CPU Load
  - TCP Segmentation Off Load and jumbo frames
- Throughput
  - How to increase?
- Monitoring
  - How to identify network performance problems
- Overview of Best Practices

# Storage Management



- LUN Queue Depth
- SAN design
- Monitoring
- Storage Response Time
- Overview of Storage Best Practices from vendor



# The Rundown



- Introduction
- Hollywood
- Virtualization
- Virtualization Capacity Management Challenges
- Virtualization Key Metrics
- The Script - Telling the story
- The End

# Needs/Benefits of Capacity Management

## Needs

To manage increasing complexity of IT

To meet growing business dependence on IT

To deliver business demand for IT

## Benefits

Better forecasts for future requirements

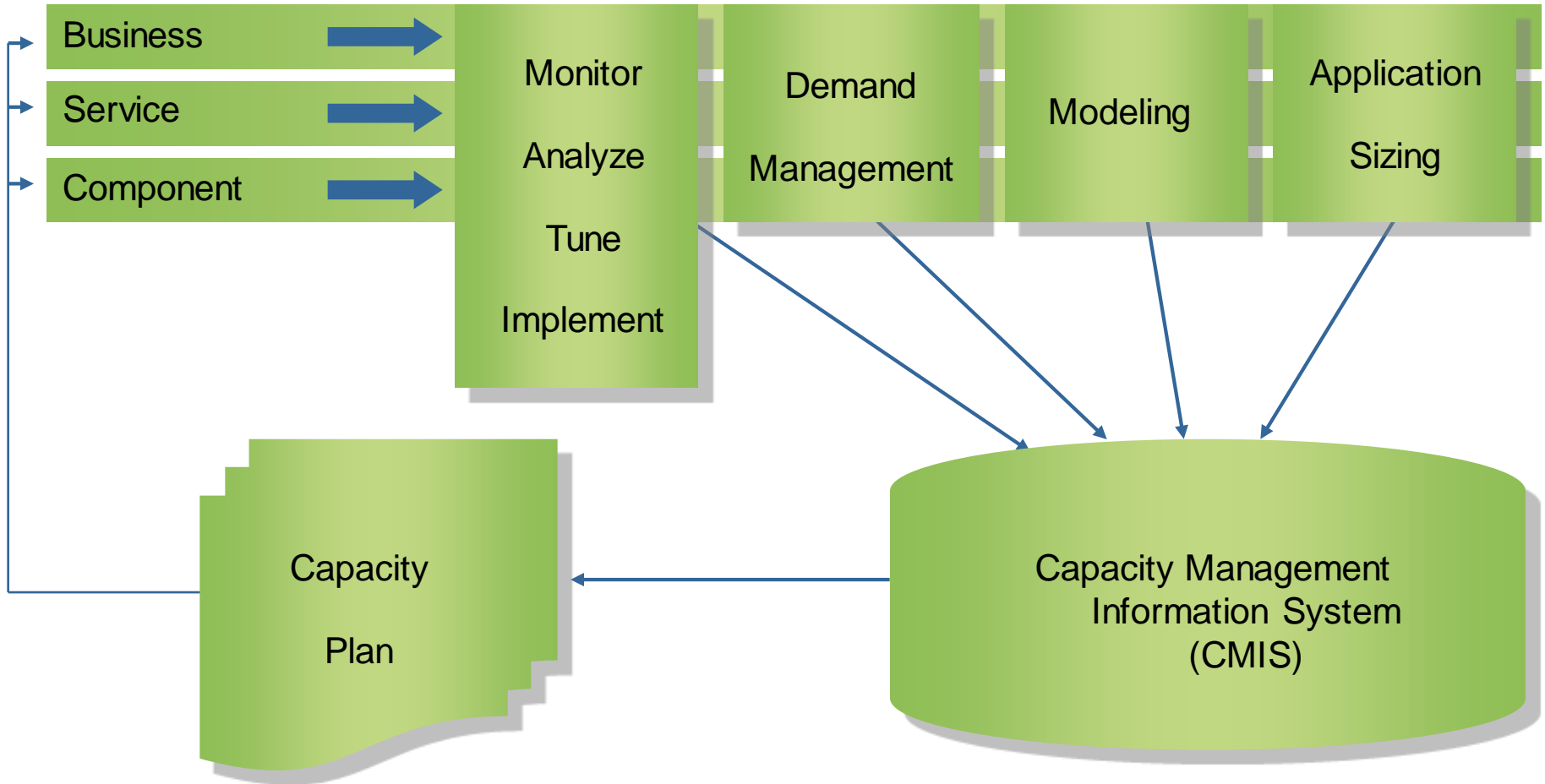
Fewer service disruptions

Fit for purpose service provision within budget

Proactive Business Availability and Capacity provision

Efficient use of resources

# ITIL Capacity Management - Activities



# Virtualization Capacity Management Challenges

- Virtual Machine Workloads
- Queuing – CPU, Memory, I/O, Network
- Sufficient CPU
- Sufficient Memory
- I/O Throughput
- Storage layout
- 32 vs. 64 bit OS and applications
- New technology usage
- Hidden overhead

# The Rundown



- Introduction
- Hollywood
- Virtualization
- Virtualization Capacity Management Challenges
- **Virtualization Key Metrics**
- The Script - Telling the story
- The End

# Understand the hierarchy

- Cluster
  - Individual infrastructure clusters
  - The wider estate
- Host
  - How is the host performing
  - How much capacity is available
  - Movement of Virtual Machines
- Guest
  - Check vCPU settings
  - Check memory settings
  - Application performance & interaction

# Key Performance Metrics



## vSphere

### CPU

Avg. CPU Usage in MHz  
CPU Ready Time

### Memory

Avg. Memory Usage in KB  
Balloon KB  
Swap Used KB  
Consumed  
Active

### I/O

Queue Latency  
Kernel Latency  
Device Latency

## Hyper-V

### CPU

Hypervisor Logical Processor %  
Hypervisor Virtual Processor %

### Memory

Dynamic Memory Balancer  
Dynamic Memory Pressure  
VM Vid Partition

### I/O

Virtual Storage Device  
Virtual Network Adapter

# Virtual Machine Capacity Management



- Selecting the correct operating system
- Virtual machine timekeeping
- SMP guidelines
- Memory considerations
- Review best practices, internal and industry standards



# Application Virtualization

Resource	Application	Category
CPU	CPU-intensive	Green (with latest HW)
	More than 8 CPUs	Red
Memory	Memory-intensive	Green (with latest HW)
	Greater than 255GB RAM	Red
Network bandwidth	1-27Gb/s	Yellow
	Greater than 27Gb/s	Red
Storage bandwidth	10-250K IOPS	Yellow
	Greater than 250K IOPS	Red

# The Rundown



- Introduction
- **Hollywood**
- Virtualization
- Virtualization Capacity Management Challenges
- Virtualization Key Metrics
- **The Script - Telling the story**
- The End

# Different presentation mediums

- Dashboards / Web portals
  - More suited for technical staff and business users
  - Different look for each area
- Word documents
  - More suited for “C” level executives
  - Upfront summary is critical
  - Historical document
- PowerPoint documents
  - Used for different types of presentations
  - Good for discussion groups
  - What is left out?

# Application Summary

Arrival 1

Arrival 2

Arrival 3

Final

Comments

**Call Center**



Infrastructure data patterns indicate spikes to be monitored and assessed in next model

**Warehouse**



I/O bottleneck; new architecture

**Data Mart**



I/O bottleneck; investigation in progress

**Sales History**

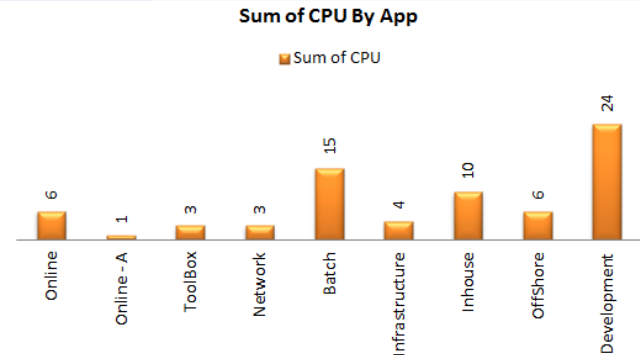
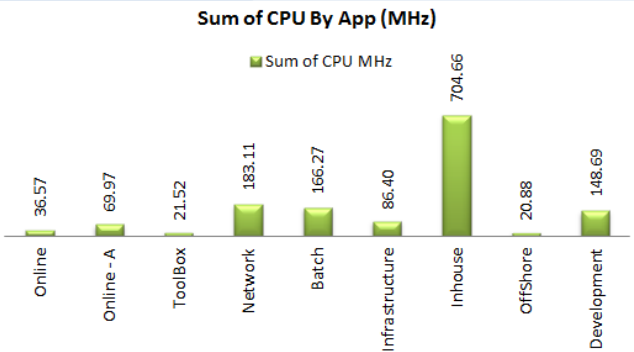
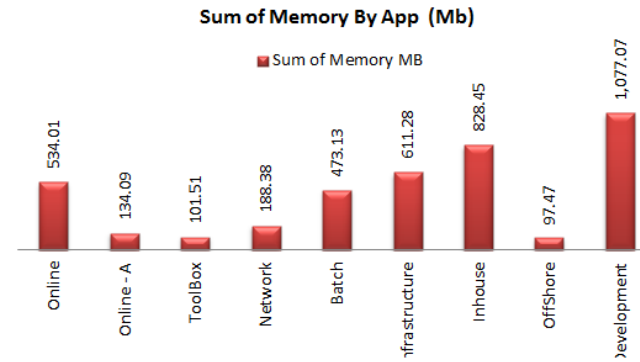
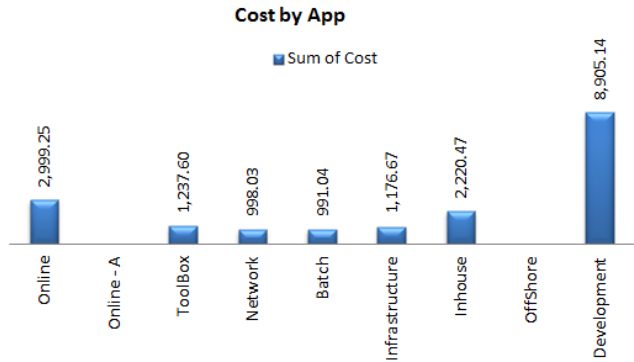



I/O bottleneck; investigation in progress

\* Scaled horizontally

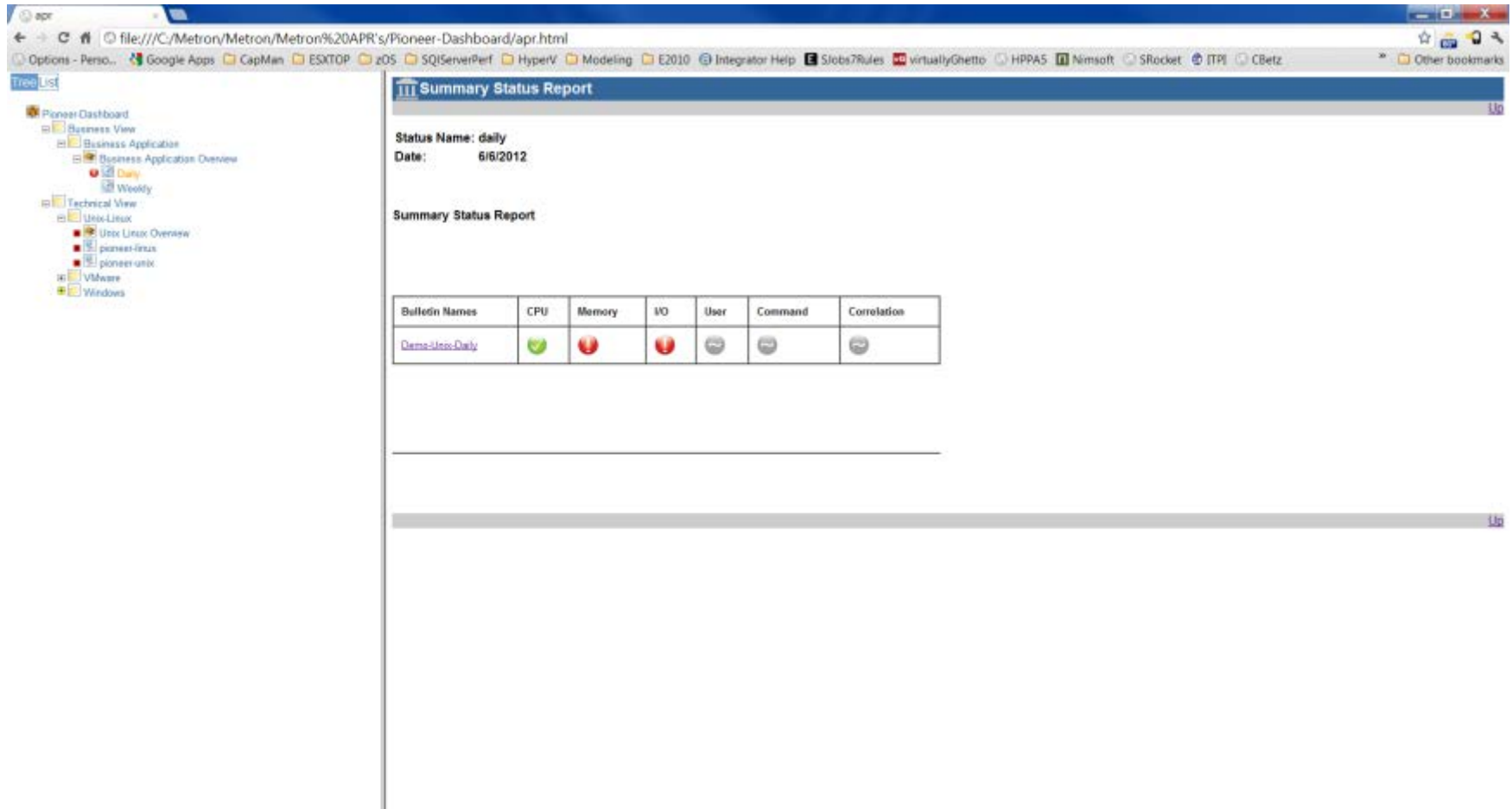
Complete your sessions evaluation online at [SHARE.org/BostonEval](https://SHARE.org/BostonEval)

# Resources and Costs



Idle - Calc Dashboard 

# Dashboard – Overview Scorecard



Summary Status Report

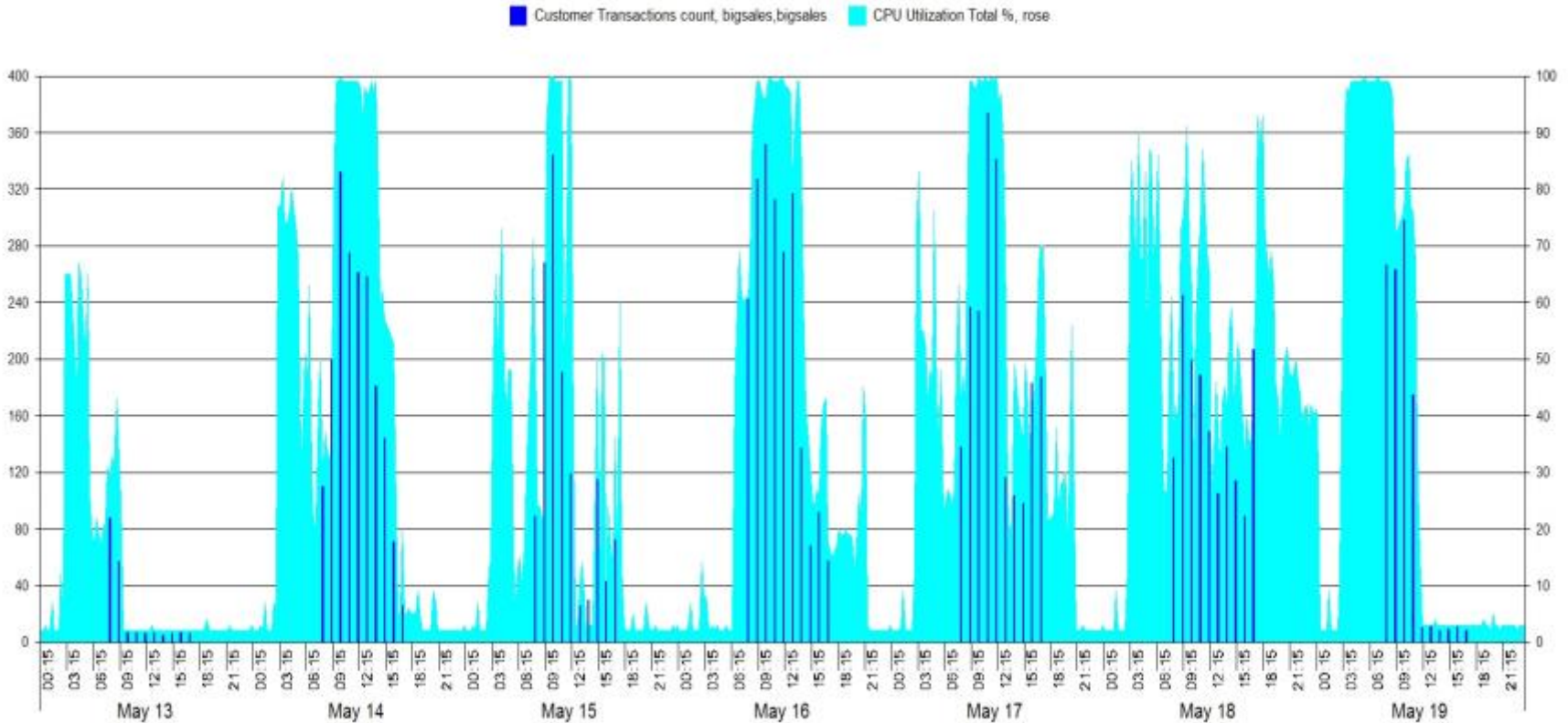
Status Name: daily  
Date: 6/6/2012

Summary Status Report

Bulletin Names	CPU	Memory	IO	User	Command	Correlation
<a href="#">Demo-Linux-Daily</a>						

# Business vs. Technical

Sales Transactions vs. CPU Utilization  
Application Server Group: Store Sales  
Reporting Period: 05/28 to 05/28

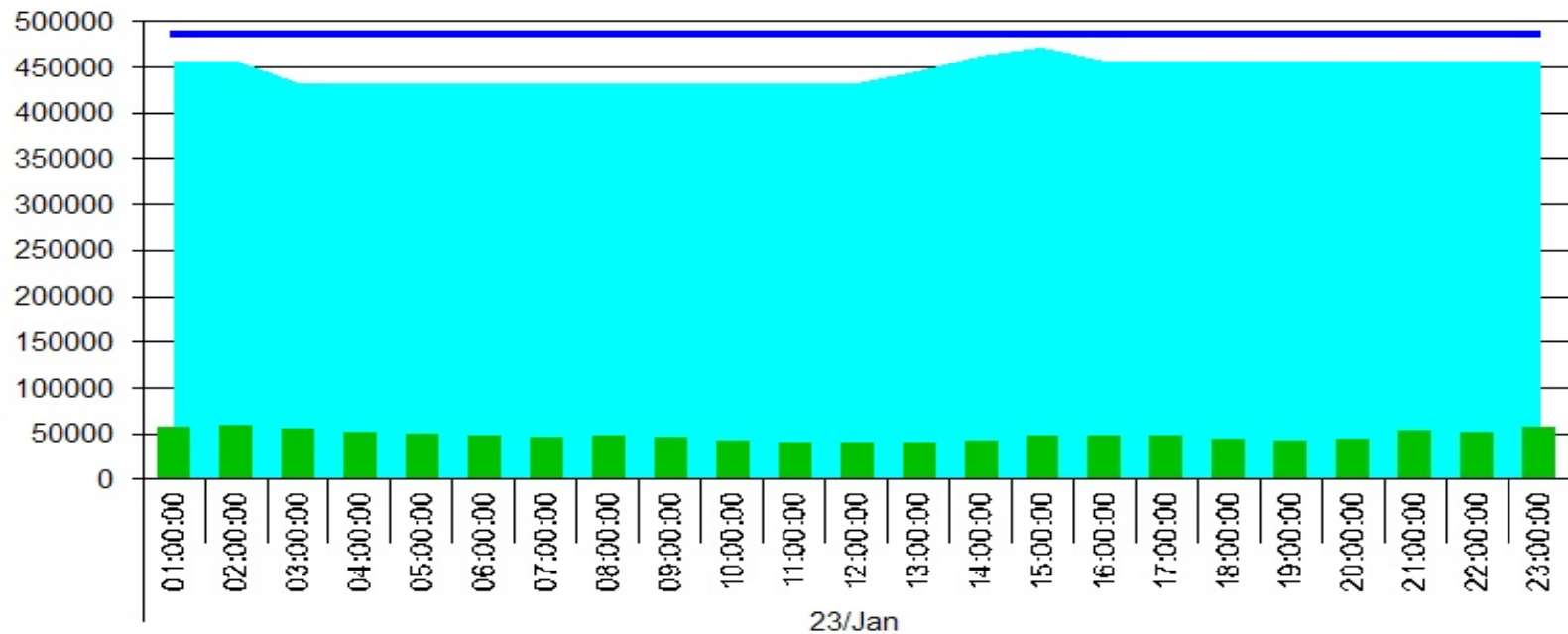


# Headroom



## vCenter1 Virtual Center Head Room CPU Analysis Period: 1/23 - 1/23

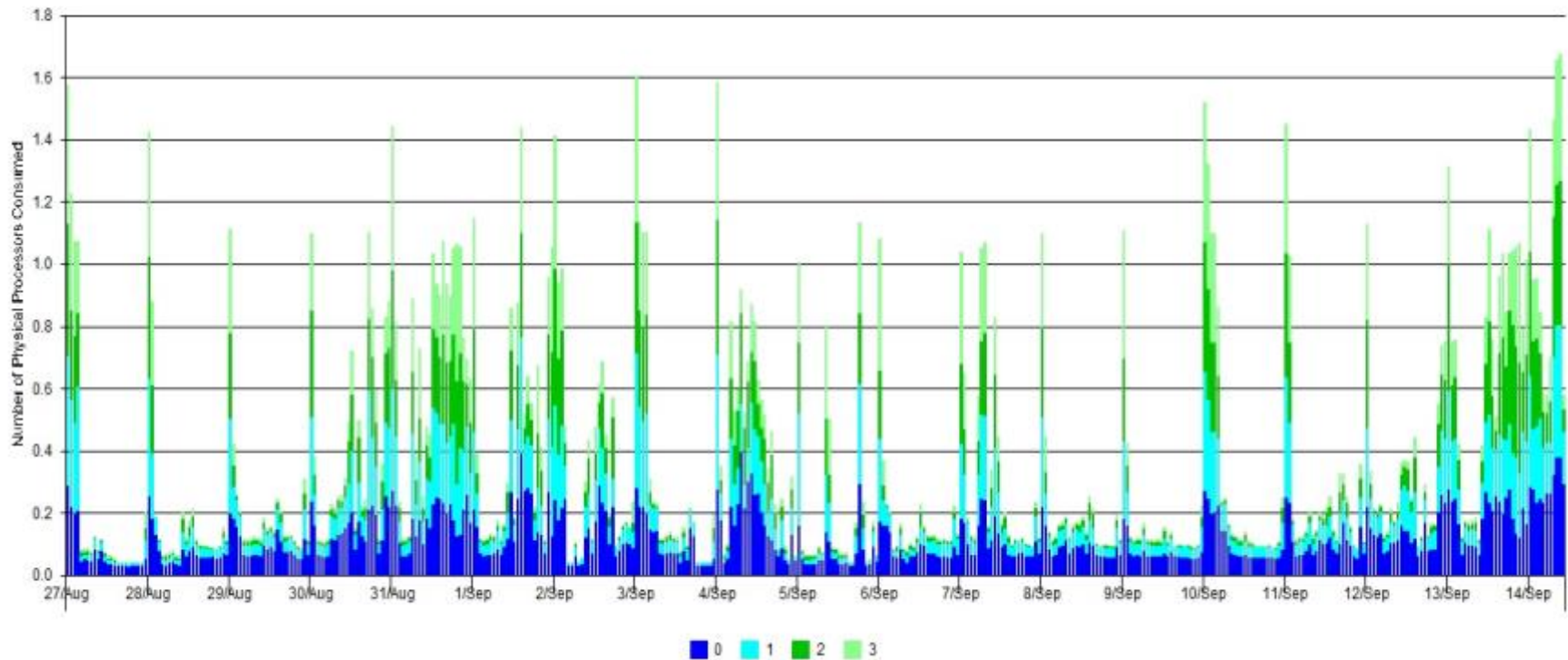
— Total Host CPU Power MHz  
■ Effective CPU Available for VMs MHz  
■ Average VM CPU Usage MHz





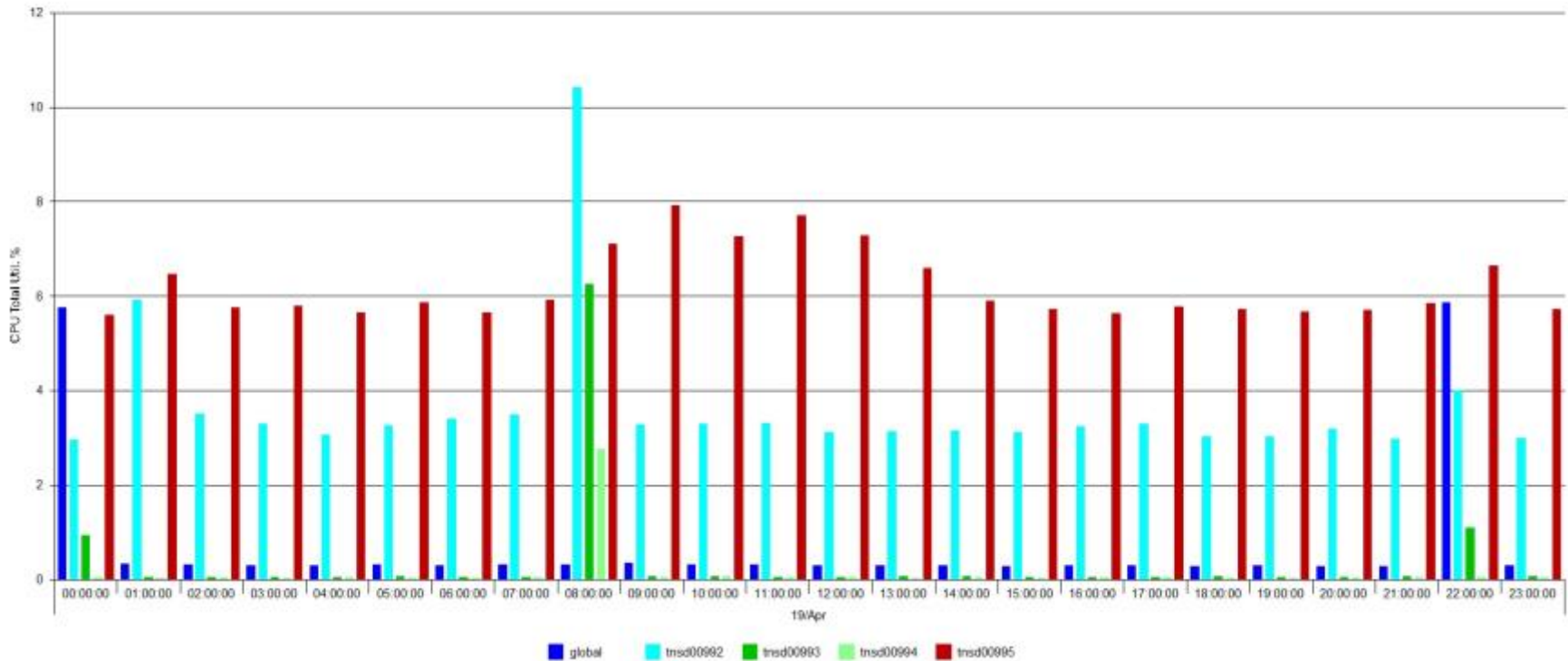
# AIX LPAR

AIX LPAR  
Processors Consumed  
Analysis Period: August through September



# Solaris Zones

Solaris Zones Aggregated CPU Total % (Daily)



# Summary



- Makes you ask questions – What is Virtualization? What is the benefit / affect on our business?
- Keeps you engaged and holds your attention – Information is clear. Reports are concise, presentation is appropriate
- A story is told and the viewer can relate to it in some way – What is happening and how much will cost?
- The ending relates to the beginning – What is the benefit / affect on our business?

# The End

