Performance Monitoring and Capacity Planning

John Paul & Chris Hayes

Session: ADC0199

VMWORLD 2006

Acknowledgements

- With special thanks and contributions from:
  - **Greg McKnight**, IBM Distinguished Engineer
    IBM Systems and Technology Group
  - **Victor Barra,** Systems Analyst
    Siemens Medical Health Solutions VMware Team
  - **Jennifer Eber,** Graphic Artist
  - And a few others who have asked to remain anonymous but have been invaluable by providing input to this presentation

Performance Monitoring and Tuning in Simple Terms

- The goal of performance monitoring in the virtual world is to determine the impediments to the full utilization of the Core Four resources (CPU, Memory, Network I/O, Disk I/O)
  - Understand the workload within a single virtual machine
  - Understand the virtualization overhead with different mixes of workloads
  - Understand the capacity of the underlying SAN and Network infrastructures, from both a cumulative as well as individual path perspective
    - Performance bottlenecks in the underlying storage and network infrastructure can affect the individual virtual machines, ESX host machines, or entire ESX farms
    - I/O and Network performance metrics may be challenging to monitor and may need special clients from the hardware vendors

Performance Monitoring and Tuning in Simple Terms

- The goal of performance tuning in the virtual world is to remove the impediments to the full utilization of the Core Four resources while minimizing the resources consumed for the virtualization overhead
  - Resolve application and configuration issues within a guest machine
  - Put complimentary workloads on the same host server
  - Design and implement the needed bandwidth for SAN and Network with particular focus on the bottlenecks of each path
    - A solid understanding of the types of I/O used and corresponding I/O architecture for network storage is now a necessary skill for the virtualization team

Performance monitoring and tuning is a continual process!

Five Contexts of Virtualization

Physical Machine

Virtual Machine

ESX Host Machine

ESX Host Farm/Cluster

ESX Host Complex

Remember the virtual context

Establish the Basic Performance Analysis Approach

- **Identify** the virtual context of the reported performance problem
- **Monitor** the performance within that virtual context for an **overview**
  - Start with the overall health of the farm/complex, looking for atypical resource consumers (individual virtual machines)
  - Analyze those virtual machines
    - Identify processes using the largest amount of the Core Four resources
    - Apply a reasonability check on the resources consumed – "Is the amount of resources consumed characteristic of this particular application or task for the server processing tier?"
    - Look for repeat offenders!
- **Expand** the performance monitoring to each virtual context as needed
  - Are other workloads influencing the virtual context of this particular application and causing a shortage of a particular resource?
- **Drill down or up** if the higher level diagnostics cannot identify the problem
- **Remedy** the problem
  - Correct the application configuration
  - Adjust the resources assigned to the virtual context
  - Remove the infrastructure problem which is degrading this virtual context

Virtual Context: Physical Machine

Monitoring Tools: Perfmon (Report View), Task Manager

Physical server resources often hide performance problems caused by less then optimum application and operating system configurations

Establish a baseline for the Core Four resource consumption and the expected demands on the underlying storage and network infrastructure

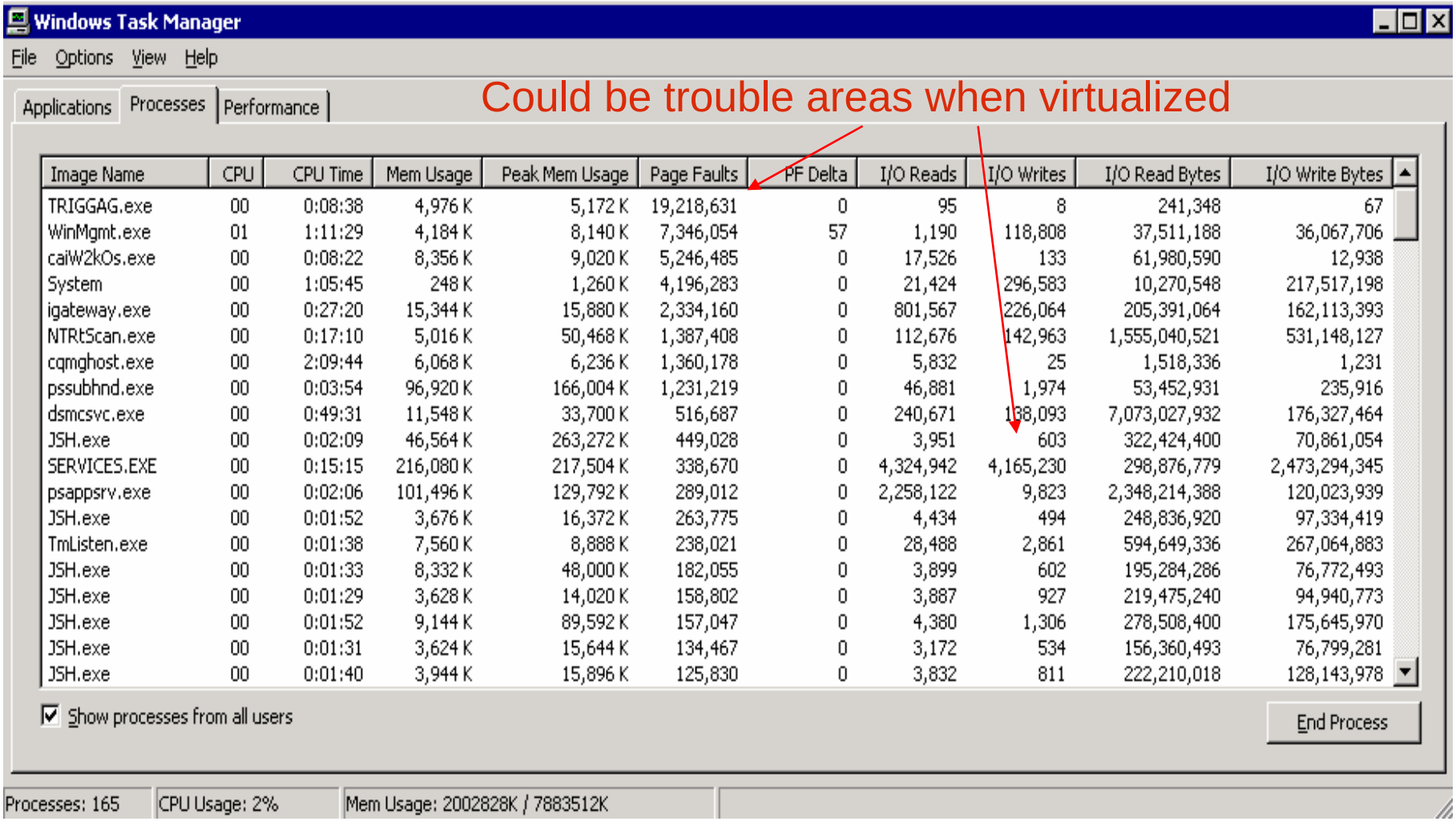
CPU
- Average physical CPU utilization
- Peak physical CPU utilization
- CPU Time
- Processor Queue Length

Memory
- Memory Usage
- Peak Memory Usage
- Page Faults
- Page Fault Delta

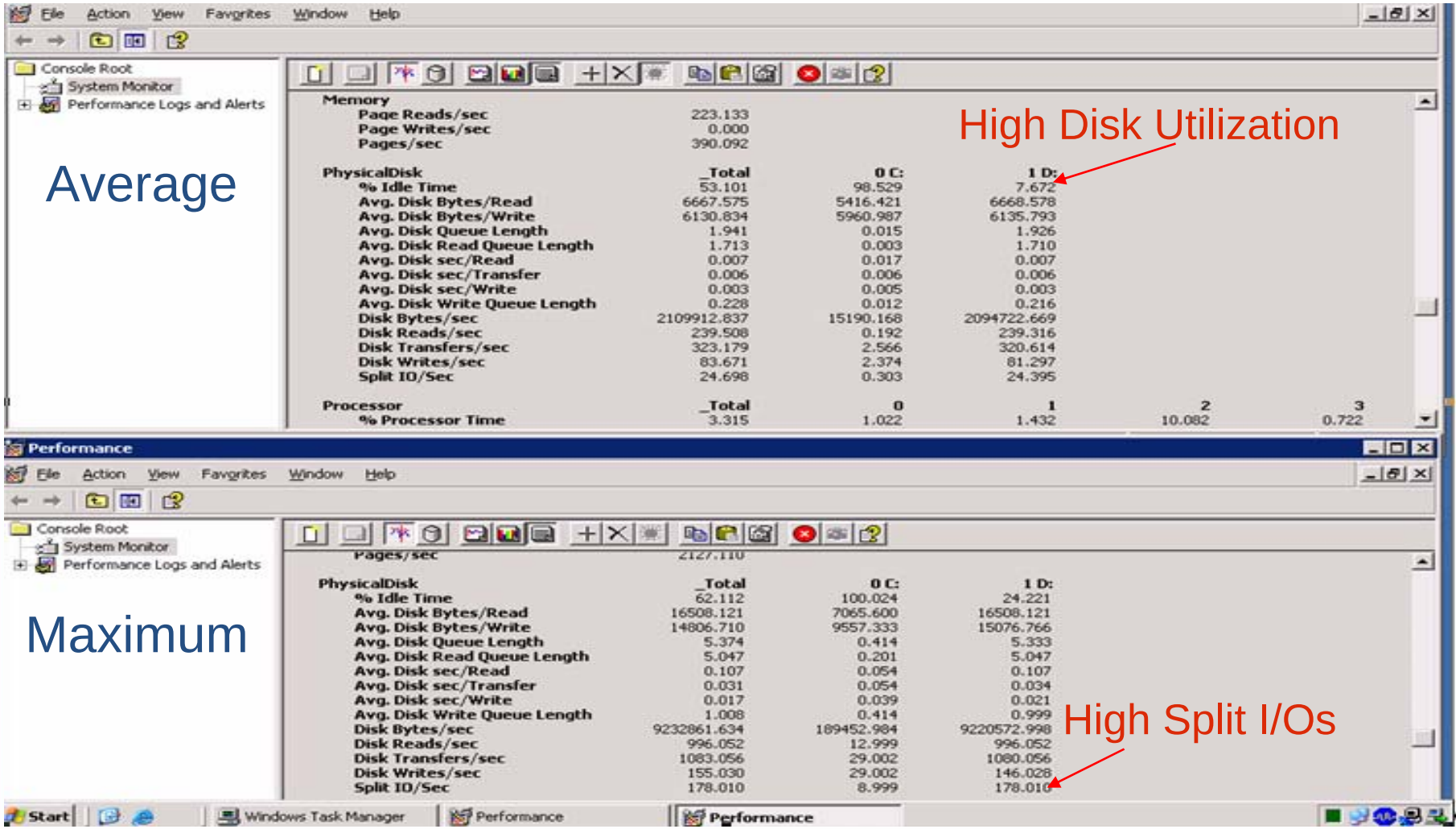
Disk
- I/O Reads
- I/O Writes
- I/O Read Bytes
- I/O Write Bytes
- Split IO/Sec
- Disk Read Queue Length
- Disk Write Queue Length
- Average Disk Sector Transfer Time

Network
- Bytes Received/second
- Bytes Sent/Second
- Output queue length

Virtual Context: Physical Machine Tools - 1



Virtual Context: Physical Machine Tools - 2



Virtual Context: ESX Host Farm/Cluster

Monitoring Tools: Virtual Center, Virtual Infrastructure Client, SAN/Network Monitors

Farm performance degradation could indicate underlying infrastructure problems in either the disk and/or network areas. Look for trends on a daily/weekly basis at this level.

VirtualCenter samples data at 1 minute intervals but averages results based upon the selected Display Period (day, week, month, year) so the longer the Display Period the longer the averaging period

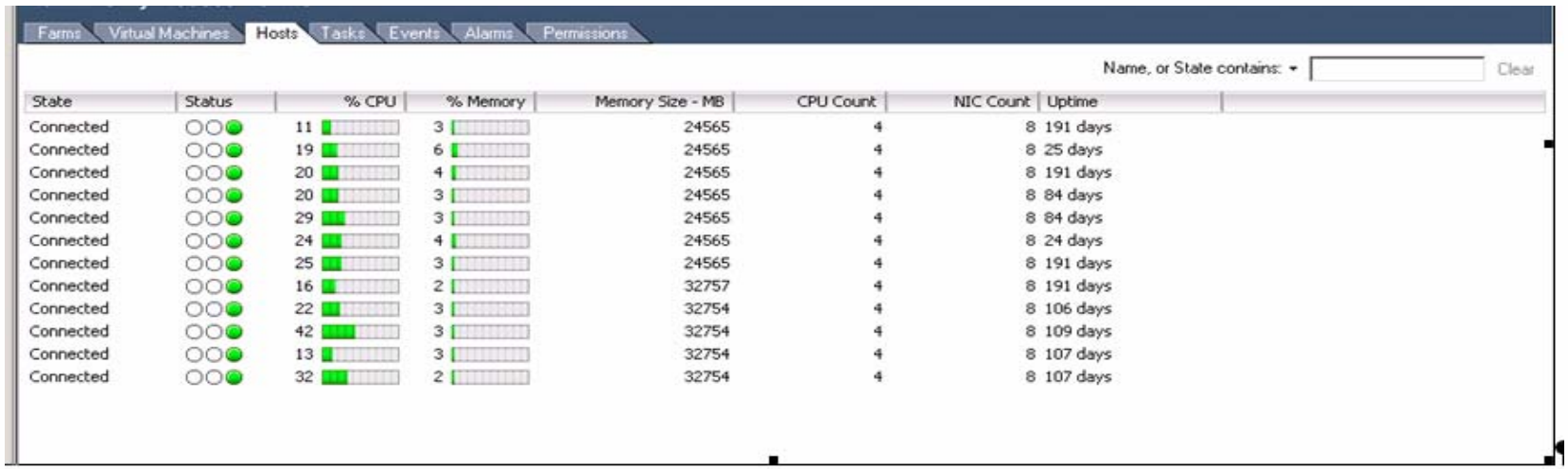
CPU
- Average physical CPU utilization
- Peak physical CPU utilization

Memory
- Average Memory Usage
- Peak Memory Usage

Disk
- I/O Reads
- I/O Writes
- I/O Read Bytes
- I/O Write Bytes
- Average Disk Sector Transfer Time
- SAN hot spots and disk utilization
- SAN cache hit ratio, based on I/O types

Network
- Bytes Received/second
- Bytes Sent/Second
- Network utilization

Virtual Context: ESX Host Farm/Cluster Tools - 1



State	Status	% CPU	% Memory	Memory Size - MB	CPU Count	NIC Count	Uptime
Connected		11	3	24565	4	8	191 days
Connected		19	6	24565	4	8	25 days
Connected		20	4	24565	4	8	191 days
Connected		20	3	24565	4	8	84 days
Connected		29	3	24565	4	8	84 days
Connected		24	4	24565	4	8	24 days
Connected		25	3	24565	4	8	191 days
Connected		16	2	32757	4	8	191 days
Connected		22	3	32754	4	8	106 days
Connected		42	3	32754	4	8	109 days
Connected		13	3	32754	4	8	107 days
Connected		32	2	32754	4	8	107 days

Virtual Context: ESX Host Farm/Cluster Tools - 2

Virtual Context: ESX Host Machine

Monitoring Tools: ESXTop, VirtualCenter, Management Console, Virtual Infrastructure Client

Hosts and Guest Machines utilize a single active path for disk I/O. Use SAN monitoring tools to diagnose hot spots and LUN queuing

Metrics to Look At:

- CPU utilization and distribution
- Physical CPU load average
- Logical CPU utilization and distribution
- CPU Effective Use
- Memory Usage
- Disk Reads/second
- Disk Writes/second
- NIC MB transmit/second
- NIC MB write/second
- % Used CPU (high consuming VMs)
- %Ready to Run
- %System (should be less then 5% total)
- %Wait
- Allocated VM memory
- Active VM memory

Virtual Context: ESX Host Machine Tools - 1

ESXTop 3.0.x

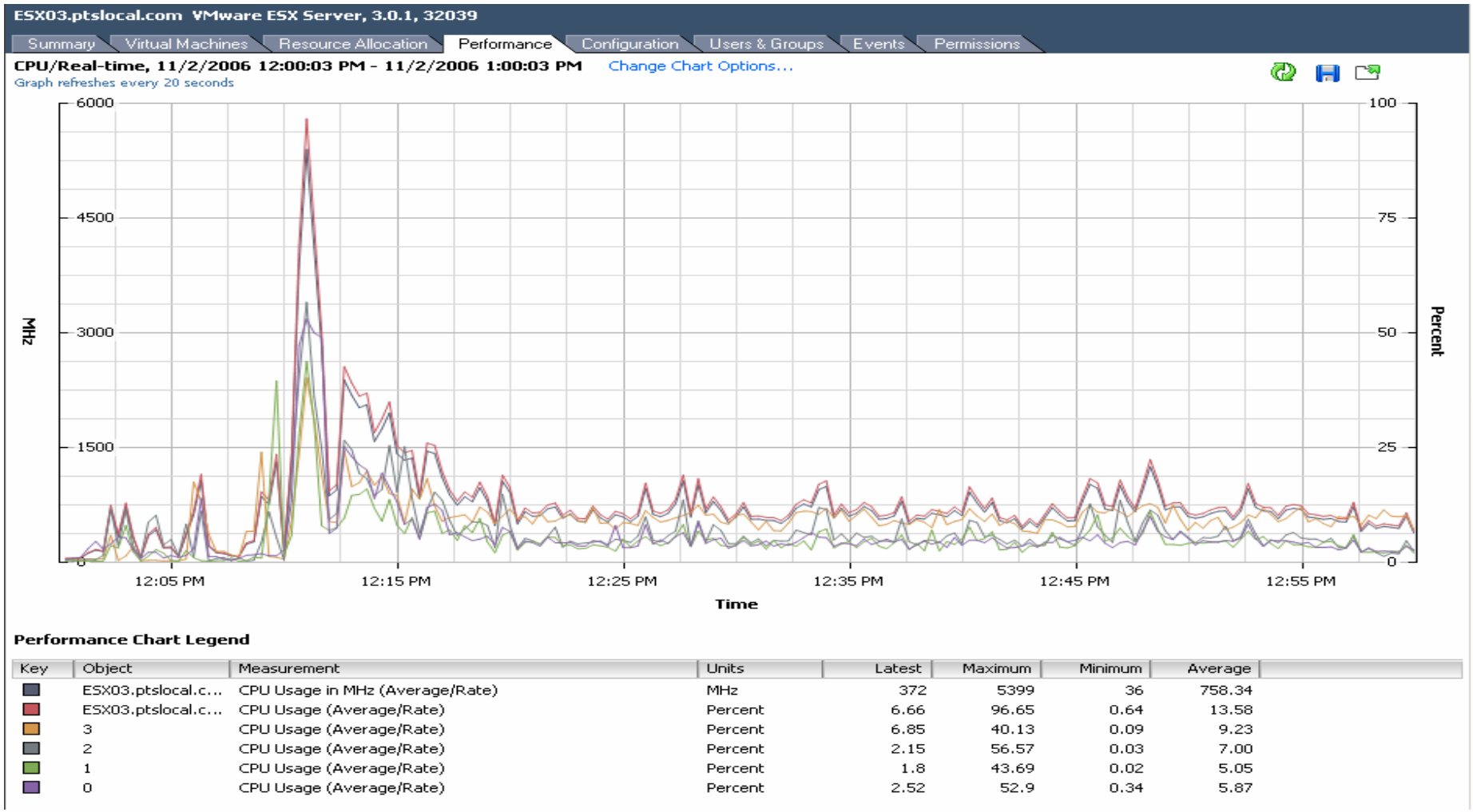
root@ESX03:~

```
 9:54:01am  up 11 days, 23:42, 62 worlds; CPU load average: 0.14, 0.14, 0.14
PCPU(%):    9.84,   11.40 ;    used total:   10.62
LCPU(%):    5.80,    4.04,    3.41,    7.99
CCPU(%):    2 us,    2 sy,   94 id,    2 wa ;          cs/sec:     113

      ID      GID NAME             NMEM     %USED      %SYS    %OVRLP      %RUN     %WAIT
       1        1 idle                4   178.83      0.00      0.00      3.22      0.00
       2        2 system              5     0.01      0.00      0.00      0.00    500.00
       6        6 console             1     5.28      0.08      0.02      5.53     29.80
       7        7 helper             13     0.01      0.00      0.00      0.02   1300.00
       8        8 drivers             7     0.00      0.00      0.00      0.01    700.00
       9        9 vmotion             1     0.00      0.00      0.00      0.00    100.00
      12       12 vmware-vmkauthd     1     0.00      0.00      0.00      0.00    100.00
      13       13 Virtual Center      5     5.63      0.01      0.08      5.60    404.46
      14       14 WIN2K3001           5     1.51      0.00      0.03      1.20    464.45
      15       15 VMWARE_WIN2K3_t     5     3.28      0.00      0.04      1.96    445.71
      16       16 WinXP_Template      5     1.33      0.00      0.04      1.32    446.64
      17       17 WinXP001            5     2.19      0.00      0.06      1.86    435.51
      18       18 WIN2K3002           5     1.83      0.00      0.03      1.43    467.92
```

ESXTop 2.5.x

```
 12:40pm  up 191 days, 15:25, 30 worlds,  load average: 0.27, 0.22, 0.21, 0.17
PCPU:   22.17%,   29.56%,   29.56%,   18.47% :   24.94% used total
LCPU:   19.71%,    2.46%,   22.17%,    7.39%,   22.17%,    7.39%,    9.85%,    8.62%
MEM: 24770560 managed(KB), 18515968 free(KB) :   25.25% used total

  VCPUID  WID WTYPE      %USED %READY     %SYS    %WAIT EMIN %EUSED     %MEM      SHRD      PRVT
     302  302 vmm       45.57   3.39     0.04    49.26   22  45.57    27.00    115.65    396.35
     136  136 idle      39.41   0.00     0.13     0.00       39.41     0.00      0.00      0.00
     135  135 idle      39.41   0.00     0.06     0.00       39.41     0.00      0.00      0.00
     134  134 idle      39.41   0.00     0.12     0.00       39.41     0.00      0.00      0.00
     131  131 idle      39.41   0.00     0.12     0.00       39.41     0.00      0.00      0.00
     133  133 idle      29.56   0.00     0.08     0.00       29.56     0.00      0.00      0.00
     132  132 idle      29.56   0.00     0.10     0.00       29.56     0.00      0.00      0.00
     130  130 idle      29.56   0.00     0.10     0.00       29.56     0.00      0.00      0.00
     129  129 idle      29.56   0.00     0.12     0.00       29.56     0.00      0.00      0.00
     383  383 vmm       17.24   1.23     0.00    83.75   22  17.24    12.00     30.39    737.61
     209  209 vmm        8.62   0.62     0.02    88.67   44   8.62    16.00    364.11    659.89
     232  232 vmm        6.47   0.42     0.00    88.67   22   6.47     4.00    329.25    694.75
```

Virtual Context: ESX Host Machine Tools - 2



Performance Chart Legend

Key	Object	Measurement	Units	Latest	Maximum	Minimum	Average
	ESX03.ptslocal.c...	CPU Usage in MHz (Average/Rate)	MHz	372	5399	36	758.34
	ESX03.ptslocal.c...	CPU Usage (Average/Rate)	Percent	6.66	96.65	0.64	13.58
	3	CPU Usage (Average/Rate)	Percent	6.85	40.13	0.09	9.23
	2	CPU Usage (Average/Rate)	Percent	2.15	56.57	0.03	7.00
	1	CPU Usage (Average/Rate)	Percent	1.8	43.69	0.02	5.05
	0	CPU Usage (Average/Rate)	Percent	2.52	52.9	0.34	5.87

Virtual Context: Virtual Machine

Monitoring Tools: Perfmon, Task Manager, Process Explorer ESXTop, VirtualCenter, Management Console, Virtual Infrastructure Client

It is very important that you look at performance from within the Guest Machine AND at the ESX host level to get the "true" view of performance

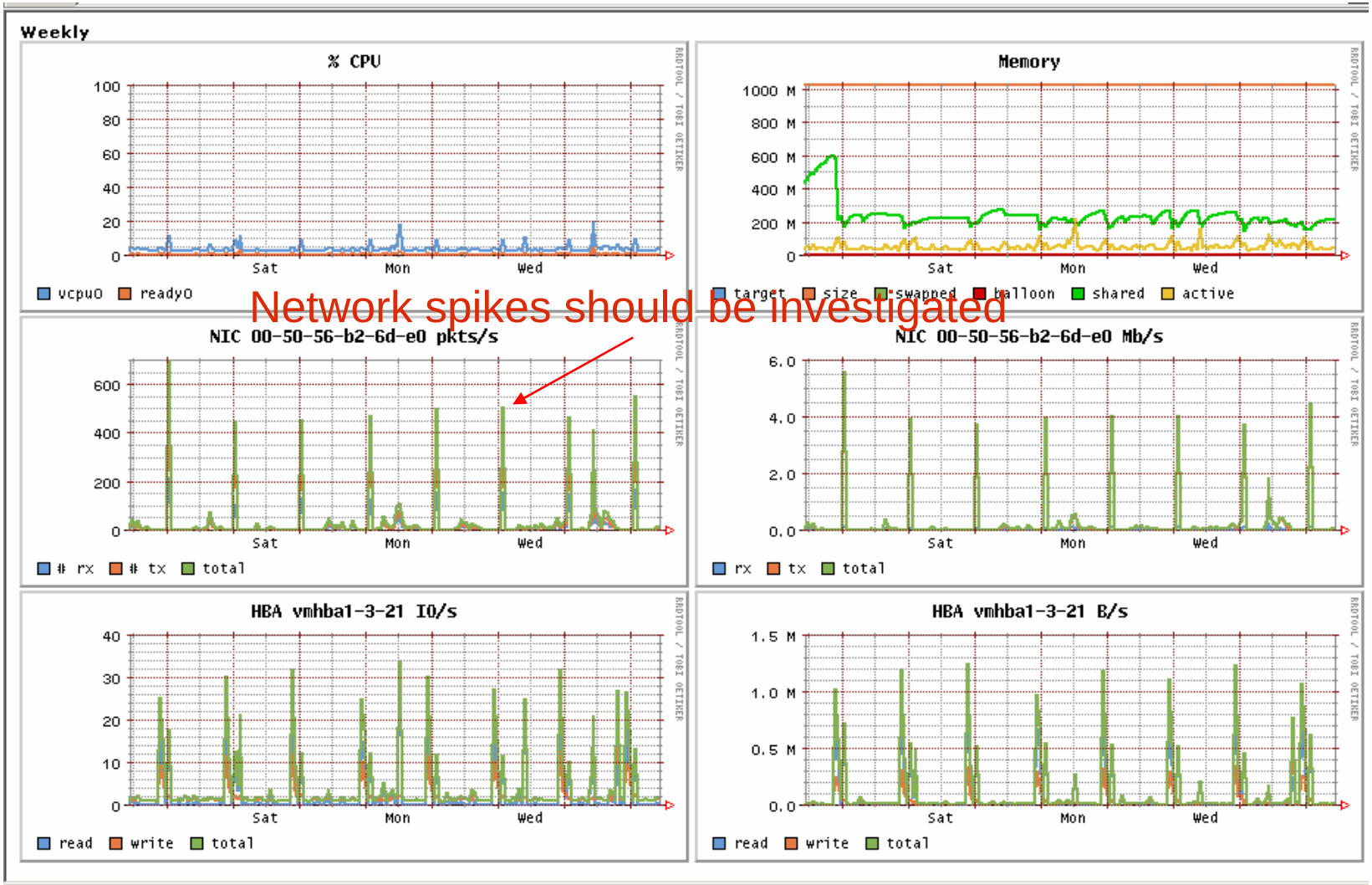
The host machine resident tools are the same as the physical machine tools, except that some of the counters are not accurate.

Additional ESX host and Virtual Center tools are used to provide a complete picture of the virtual machine

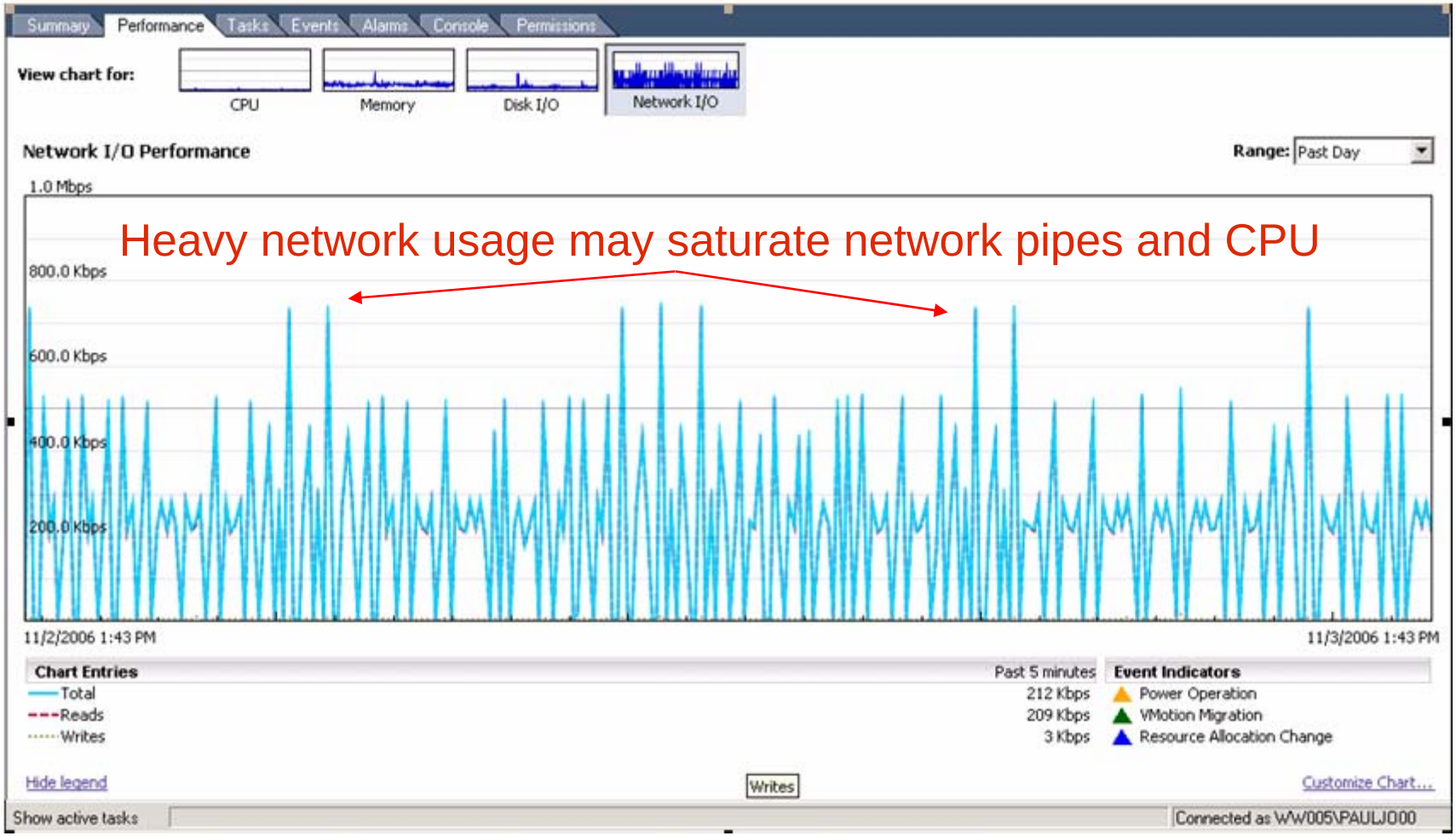
Metrics to Look At:

- Average physical CPU utilization
- Peak physical CPU utilization
- CPU Time
- Processor Queue Length
- Memory Usage
- Peak Memory Usage
- Page Faults
- Page Fault Delta
- I/O Reads
- I/O Writes
- I/O Read Bytes
- I/O Write Bytes
- Split IO/Sec
- Disk Read Queue Length
- Disk Write Queue Length
- Average Disk Sector Transfer Time
- % Used CPU
- %Ready to Run
- %Wait
- Allocated VM memory
- Active VM memory

Virtual Context: Virtual Machine Tools - 1



Virtual Context: Virtual Machine Tools - 2



Performance “Warning Signs”

- Physical CPU
  - Sustained usage of >80%
  - Unbalanced usage across processors/hyper-threads over time
  - Processor queue length per CPU>10
- Memory
  - Total paging greater than 200-300 I/O per second
- I/O
  - Most common area for performance issues
  - >20ms average sec/transfer time for physical disk
  - >3 average queue length
  - Split I/O average >1% of total disk I/O
- Network (NIC)
  - Network queueing regularly occurring
- Network (LAN/WAN)
  - Network sniffer most effective at determining usage and bandwidth

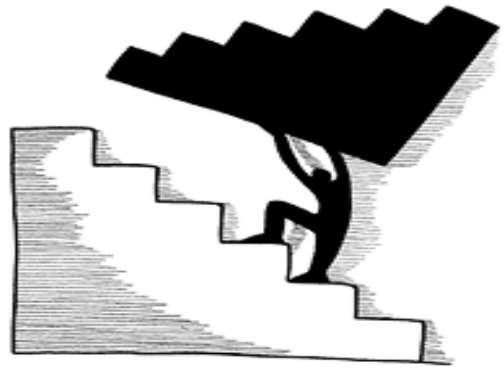
Disk Subsystem Performance Overview

- Don't just consider capacity, consider adding more disks
  - For multi-threaded I/O intensive applications, more disks = more performance
  - Random read/write workloads usually require lots of disks to scale
  - For random write intensive environments:
    - RAID-10 about 50% greater throughput than RAID-5
    - Magnitude of gain depends upon % of write commands
    - 50% slower for RAID5 (67% reads 33% writes) - typical commercial workload
- RAID Ratio of performance for comparing RAID strategies:
  - %Reads * (Physical Read Ops) + %Writes * (Physical Write Ops)
- RAID-10, RAID-1, RAID 0+1, RAID-1+0
  - Two physical disk writes per logical write request are required
  - I/O Performance = % Read * (1) + % Write * (2)
- RAID-5
  - Four physical disk I/O operations per logical random write request are required (two reads and two writes)
  - I/O Performance = % Read * (1) + % Write * (4)

When a Disk is not a Disk – SAN Considerations

- Local disk solutions typically involve an on-board or plug-in SCSI controller with a small amount of read/write cache
  - Data drives compete with system drives and paging
- SAN solutions can include network fabric, network switches, network adaptors, host bus adaptors, frame adaptors, front-end processors, microcode, a variety of bus structures, and GBs of cache
  - SAN performance analysis starts with the host machine
    - Start with disk busy, average sector transfer time, IOPS
  - At the SAN level, start with the “back end” physical disks, using SAN management tools
    - The bigger the performance problem the more likely it is in the back end disk area
    - Don’t expect much more than 100 IOPS from a physical disk
  - Work your way upwards inside the SAN, working your way to the SAN fabric
  - Remember that the SAN has the similar challenges to ESX, which is competition for shared resources
    - Look for competition at the physical disk and LUN levels
  - Random reads, random writes, sequential reads, sequential writes may get homogenized in a SAN
  - I/O block sizes can be changed as the data is moved down the I/O path
  - Native SAN tools tend to measure at larger sampling intervals so results will be smoothed
  - Though individual components of a SAN or NAS have absolute throughput limits the aggregate SAN throughput limits are not the sum of its parts

Capacity Planning Overview



Key steps for Capacity Planning

1. Identify standard Virtual Machine (guest) profile
2. Identify standard Host Server profile
3. Identify standard Storage Types and Sizing Metrics
4. Develop Critical Resource Thresholds
5. Understand Resource Replenishment Timeframes
6. Evaluate Demand, Current and Forecasted (where possible)
7. Establish Capacity Replenishment Triggers

Discussion Assumptions

- Enterprise class implementation\environment (Virtual Center VMotion, Redundancy, etc.)
- Supported ESX Standards (Hosts per Virtual Center, LUNs per Host, Shared LUNs, etc.)
- Network storage (rather than local disk)
- 4 Processing units per host (single core or multi core)
- Architect\Network Administrator Perspective
- Not a Cost Model or Chargeback Discussion!

Standard Virtual Machine Profile

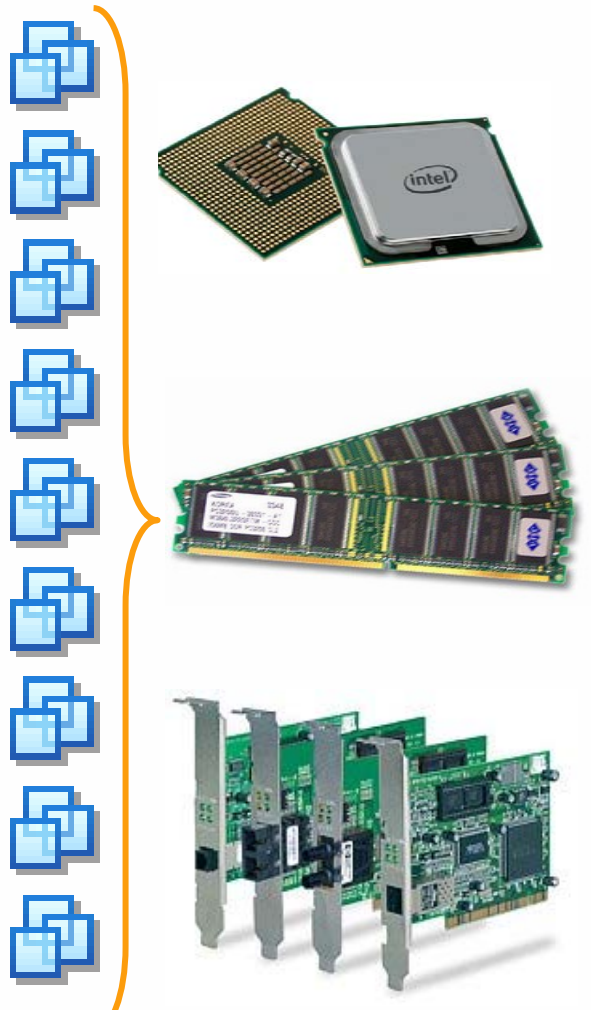
Virtual Machine Composition Elements

- CPUs\Processing Shares
- RAM
- Networking
- I\O Requirements & Needs
- Total Image Size (all .vmdk files for per image)
- Deviations Complicate the plan!

VM Profile Example (hypothetical)

- 1 Proc, 1000 Processing shares
- 1024 RAM
- 1 NIC (1 guest network path standard)
- 1.0 MBps I\O throughput
- 35 GB total image size
- Limit the Severity & Frequency of Deviations!

Develop a Baseline Standard!



Standard Host Server Profile

- **Host Server Composition Elements**
  - Processing Capability
  - Memory Capability
  - Networking Connectivity
  - Storage Connectivity
  - Build Configuration

- **Host Server Example (hypothetical)**
  - 4 Processing Units (single or multi-core)
  - 24 GB RAM
  - GB NICs: 1 GB VMotion, 1 GB Service Console, 2-6 GB dedicated to VM's (assuring redundancy & throughput)
  - 2, 2GB HBA cards (or equivalent storage connectivity)
  - Fast, Redundant storage (local) for ESX Host software

Develop a Baseline Standard!

Standard Storage Profile

- **Storage Considerations**
  - Storage Options: **ESX 2.5**--SAN; **ESX 3.0**--SAN, NAS, ISCSI, etc.
  - Manufacturer\Model\Class: Site-dependent, but *tiered* storage has benefits
  - VM I\O Requirements may predicate storage type (and *virtualization applicability*)
  - Sizing Standards impact *capacity*, *performance* and *manageability*
- LUN Sizing Example (Hypothetical)
  - LUN Size: 280GB
  - VM Size: 35GB

Find your *Sweet* Spot!

Develop a Baseline Standard!

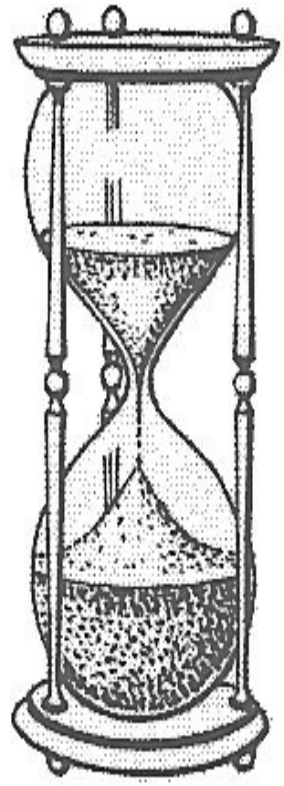
Capacity and Replenishment Planning

- **Key Capacity Planning and Replenishment Concepts Review**
  - Develop Accurate Profiles for VMs, Hosts and Storage
  - Account for the Core 4 Resources: Processing, Memory, Storage & Networking
  - Establish Critical Resource Thresholds (Platform and Organization)
  - View the *Forest* AND the *Trees:* Plan with the Farm and VMs in mind!
  - Anticipate Exceptions: Over-Consuming &\or Under-Performing Resources
  - Evaluate &\or Forecast farm Demand (resource and new provisions)
  - Develop Realistic Resource Replenishment Timeframes
  - Incorporate Tolerances to Accommodate for Delays
  - Establish Accurate, Useful Replenishment Triggers

Plan Capacity Replenishment Conservatively!

Replenishment Planning: Hosts

- **Critical Host Metrics**
  - Run Rate (VMs per processing unit)
  - Avg. Processing allocation per Farm (Peak & Non-Peak)
  - Avg. Memory allocation per Farm (Peak & Non-Peak)
  - Exception Cases per VM\Host (Proc., RAM, NIC, etc.)
  - Fulfillment Timeframes (Server, Proc., RAM, NIC, etc.)
- **Example Host Replenishment Thresholds (hypothetical)**
  - Run Rate= **>3 - 4 VMs per processing unit avg.*** (12-16\host)
  - Avg. Processing allocation per Farm (Non-Peak) @ **50-60%***
  - Avg. Processing allocation per Farm (Peak) @ **70-80%***
  - Avg. Memory allocation per Farm (Non-Peak) @ **65-70%***
  - Avg. Memory allocation per Farm (Peak) @ **75-80%***
  - Exception Case Variables: Monitor Consumption\Utilization



* Values reflect tolerance for Platform Limits, Fulfillment Timeframes and Demand

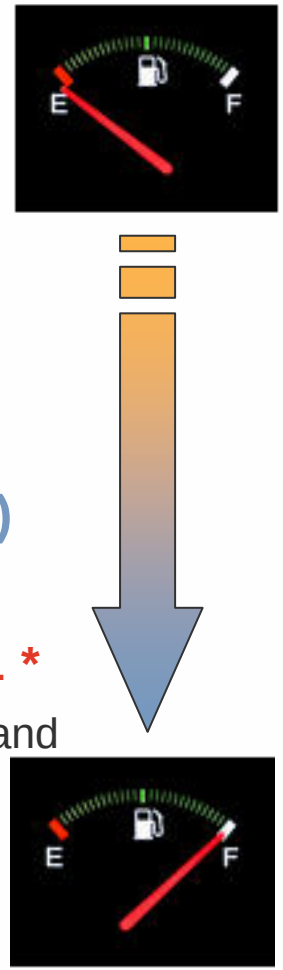
Know Your Replenishment Timeframes!

Replenishment Planning: Storage

- **Critical Storage Metrics**
  - Avg. Total utilization of available storage volumes
  - Avg. Total utilization of storage infrastructure (frame, fabric & FA)
  - Exception Cases per VM\LUN (Increased Size & I\O Requirements)
  - Fulfillment Timeframes (Volumes and Devices)

- **Example Storage Replenishment Thresholds (hypothetical)**
  - Volume Allocation\Utilization @ **80%** farm storage volume utl.*
  - Storage Infrastructure Allocation\Utilization @ **60%** infrastructure utl. *

* Values reflect tolerance for Platform Limits, Fulfillment Timeframes and Demand



Develop a Baseline Standard!

Replenishment Thresholds and Triggers Considerations

- Plot Critical Thresholds based on **Performance**, **Availability**, **Supportability**
  - CPU: VM Performance Degradation @ 82% util.
  - Memory: VM Performance Degradation @ 75% util. (avoid swapping!)
  - SAN: VM Performance Degradation @ 79-89% util. of volumes and device.
  - Run Rate: weigh supportability and fiscal pressures

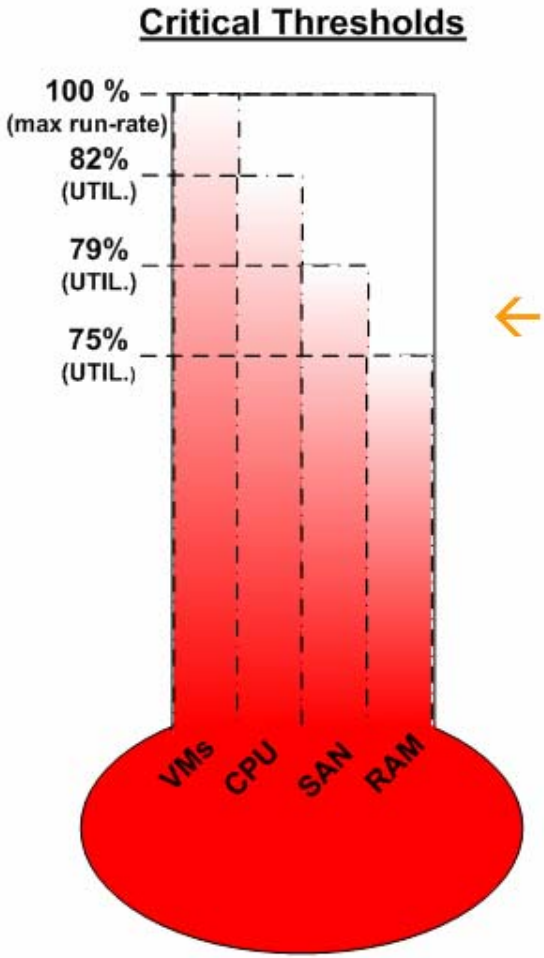
- Evaluate Replenishment Timeframes for…
  - Hosts
  - Memory
  - Storage

When?
How Much?
How Long?

Capacity Planning and The Replenishment Process

1. Determine Standard Profiles (VM, Host & Storage)
2. Understand Critical Thresholds (Platform and Organizational)
3. Evaluate Fulfillment Timeframes (Host, Memory, Storage, etc.)
4. Determine Demand: Current & Forecasted *(good luck!)*
5. Set Replenishment Triggers For Each Resource
6. Establish a Repeatable Capacity Evaluation and Replenishment Process

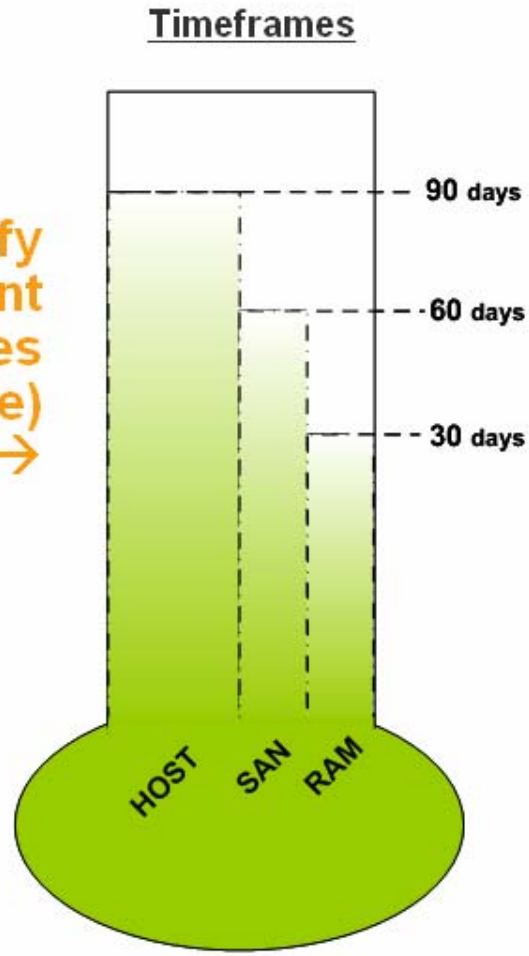
Capacity Planning and The Replenishment Process



← Understand Critical Thresholds
(Platform & Organizational)

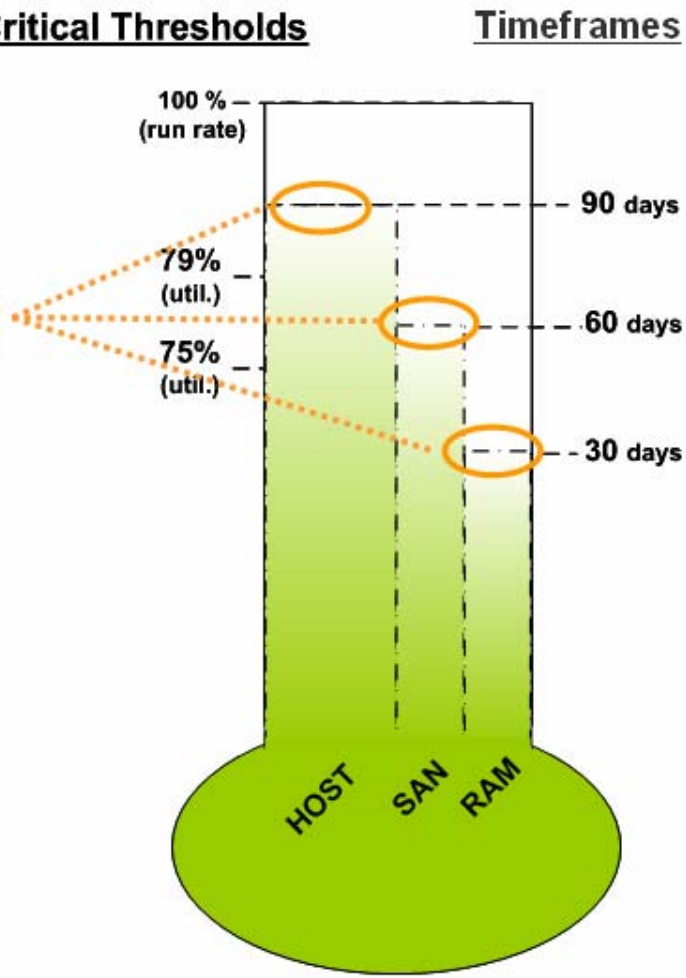
Capacity Planning and The Replenishment Process

Identify Replenishment Timeframes (worst case) →



Capacity Planning and The Replenishment Process

Set Replenishment Triggers

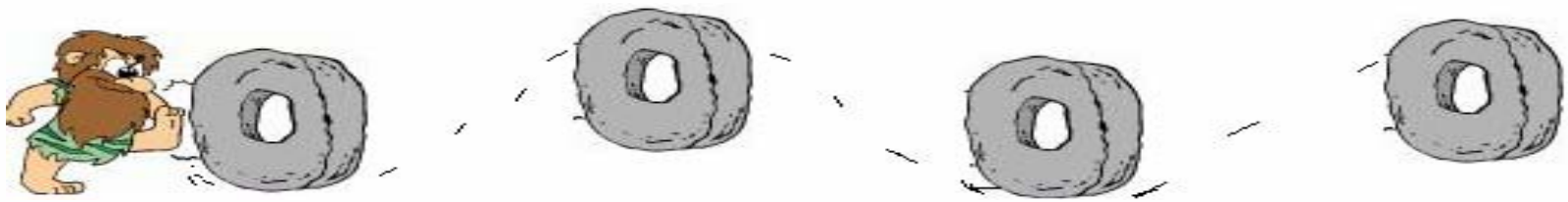


Capacity Planning: In The Eyes of the Beholder…

- **Technical\Tactical vs. Financial:**
  - How far do you push your thresholds?
  - Is everyone committed to the Capacity Plan?
- **Innovation & Rapid Adoption: opportunity for the CTO, challenge for the CFO**
  - CTO: *If we build it and they come, great…Let's build more!*
  - CFO: *Innovation is great, but will the budget support the growth?*
- **Capacity Planning does not replace a technical roadmap**
  - Technical Roadmaps provide strategic direction for the enterprise
  - Capacity Plans assume that strategic direction has been established
- **Capacity Plans and Chargeback Models live separate lives but cross paths**
  - *Capacity Plan says we need more…Do Chargeback metrics support replenishment?*
  - *There's a new version available…time to update the Capacity Plan and Chargeback Model?*
- **Budget Planning ≠ Capacity Planning!**
  - Budget Plans focus on $$; Capacity Plans focus on Profiles, Thresholds & Replenishment
  - Budgets *typically* change quarterly/yearly; Capacity Plans *typically* change with versions

Lessons Learned…

- **Innovation Adrenaline…fight the urge:** Develop a capacity plan first!
- **Growth will happen faster then you think:** Expect the unexpected!
- **Virtualization density challenges conventional thinking:** Anticipate a learning curve for the hardware, networking and storage teams
- **Master the Cost Model:** Standardize the Cost Model before deploying VMs!
- **Set the Ground Rules from the Start:** Define usage, tools & permissions early!



More Lessons Learned…

- **Plan for Capacity rather than Reacting to Capacity:** Develop a thorough Capacity Plan before you need it, not when you need it!
- **Not Because You Can…Because You Should:** With the power of virtualization comes responsibility; plan your farm with the enterprise in mind
- **Know Your Load:** Evaluate applications before virtualizing…will it fit?
- **Virtualization Changes Business:** Ramp-up before the Tidal Wave
- **Get Everyone Involved:** Collaborate with reps from Hardware, Networking, Infrastructure, Development, Project Planning, Asset Mgmt., Venders, etc.

Performance Monitoring & Capacity Planning

John Paul – johnathan.paul@siemens.com

Chris Hayes – christopher.hayes@siemens.com

Session: ADC0199

Questions?

Presentation Download

Please remember to complete your
session evaluation form
and return it to the room monitors
as you exit the session

The presentation for this session can be downloaded at
http://www.vmware.com/vmtn/vmworld/sessions/

Enter the following to download (case-sensitive):

Username: cbv_rep
Password: cbvfor9v9r

Some or all of the features in this document may be representative of feature areas under development. Feature commitments must not be included in contracts, purchase orders, or sales agreements of any kind. Technical feasibility and market demand will affect final delivery.

VMWORLD 2006