IBM Software

Managing and Operating z/VM and Linux

Jim Newell, Tivoli zSeries Automation
jnewell@us.ibm.com

Mike Sine, Advanced Technical Support, Americas
sine@us.ibm.com

October 2009

Acknowledgement of material sources

- **Bill Bitner**
- **Brian Wade**
- **Bob Neill**
- **Bill Maddox**
- **Steve Gracin**
- **Jon vonWolfersdorf**
- **Kevin Yash**
- **Mike Sine**
- **Richard Lewis**
- **John Schnitzler**

AGENDA

- **Introduction**
- **Best Practices**
- **OMEGAMON/ITM "Quick View"**
  - Integrated Monitoring Approach
  - Use Cases
- **Why IBM**
- **Bringing it all together**

Virtual Linux servers have unique challenges versus running on physical machines.

- z/VM System Programmers and Linux Administrators may not be in the organization.

- We find that it is easy to over allocate resources; therefore, our monitoring examines resource usage of hardware, hypervisor, as well as the virtual machine. Real-time and historical metrics demonstrate peaks periods as well as average runtimes.

"Best Practices"

- z/VM
  - System Scope items
    - Maintenance, Memory, Paging, DASD, MDISK Cache, VDISK, Processors/LPAR, System Utilization, DASD I/O, Spin Locks, SRM and Quick Dispatch, CP-Owned Allocation
    - Workloads: Virtual Processors, Paging

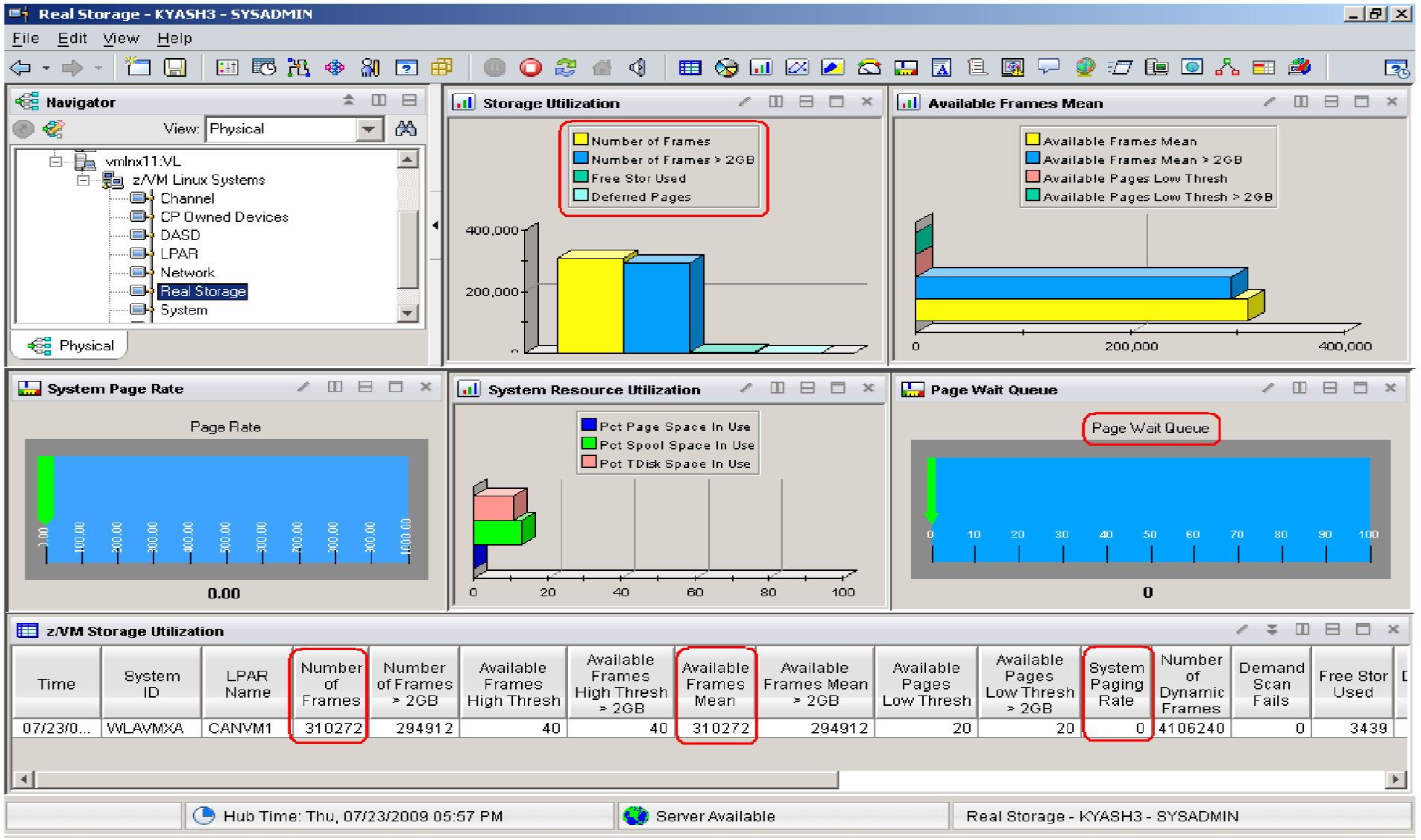
Maintenance Levels

- **Recommend maintaining current service levels.**
- **Apply latest Recommended Service Upgrade (RSU):**
  - z/VM Family
    - Released every 3-6 months
    - Contains cumulative service including all pre- and co-requisites in a pre-built format.
  - Includes service for all integrated components and the following pre-installed program products:
    - DirMaint, VM/RACF, Performance ToolKit
      - Available on tape, DVD, or electronically.
  - Separate Maintenance Stream for OMEGAMON
    - Available electronically, or via platform appropriate media.

Memory Configuration

- **Plan on a virtual to real (V:R) memory ratio in the range of 1.5:1 to 3:1.**
- **Recommend configuring some processor memory as expanded storage:**
  - Serves as high speed cache.
  - Increases consistency of response time.
  - See http://www.vm.ibm.com/perf/tips/storconf.html for the gory details.
- **Rule of Thumb - start with 25% of memory configured as expanded:**
  - Typically 2–4GB of expanded storage is sufficient, 1GB minimum.
  - The lower the paging rate, the lower the amount of expanded storage required.
  - The greater the number of page frames available in central storage above 2GB, the higher the amount of expanded storage required.

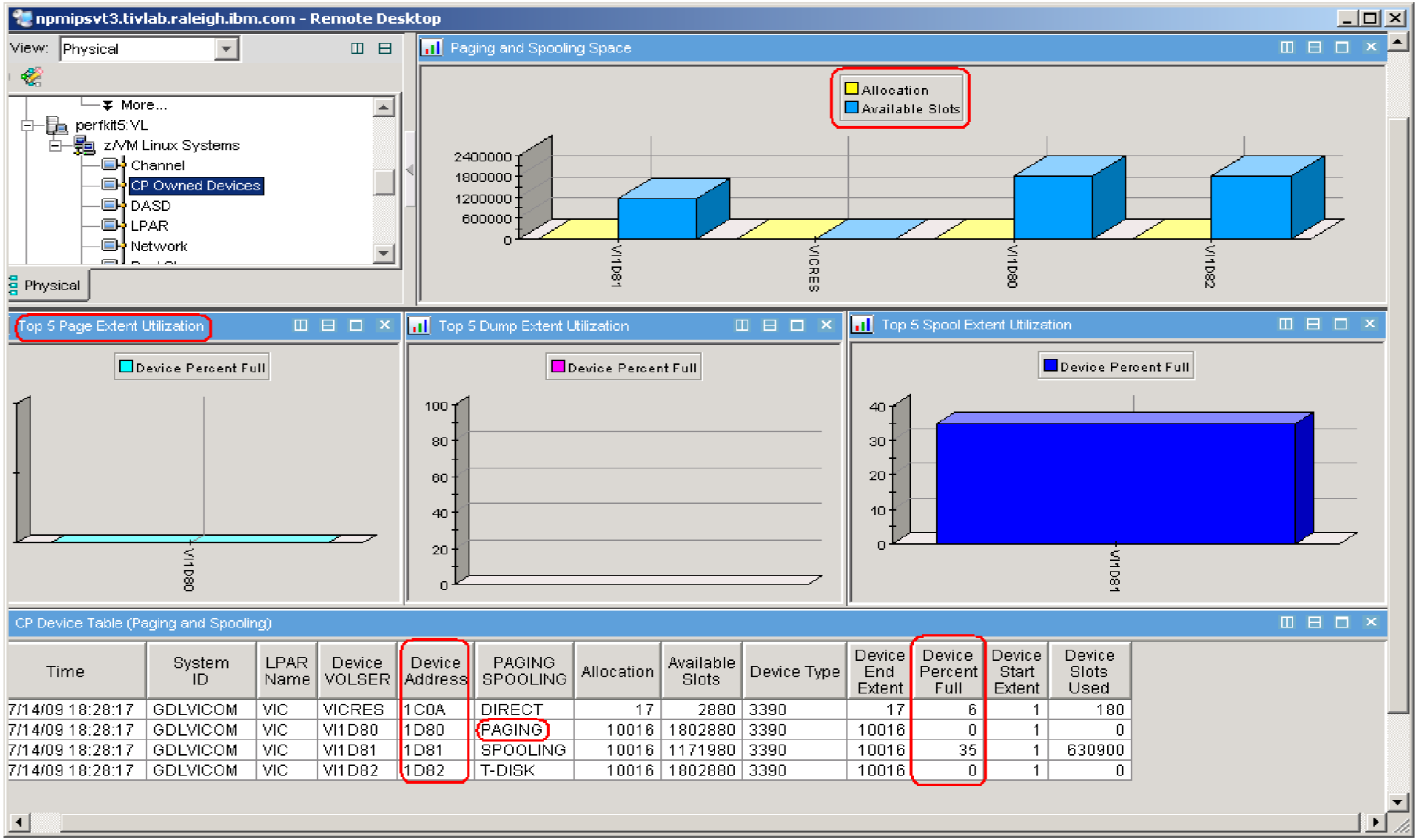
OMEGAMON Memory Configuration



Paging Subsystem

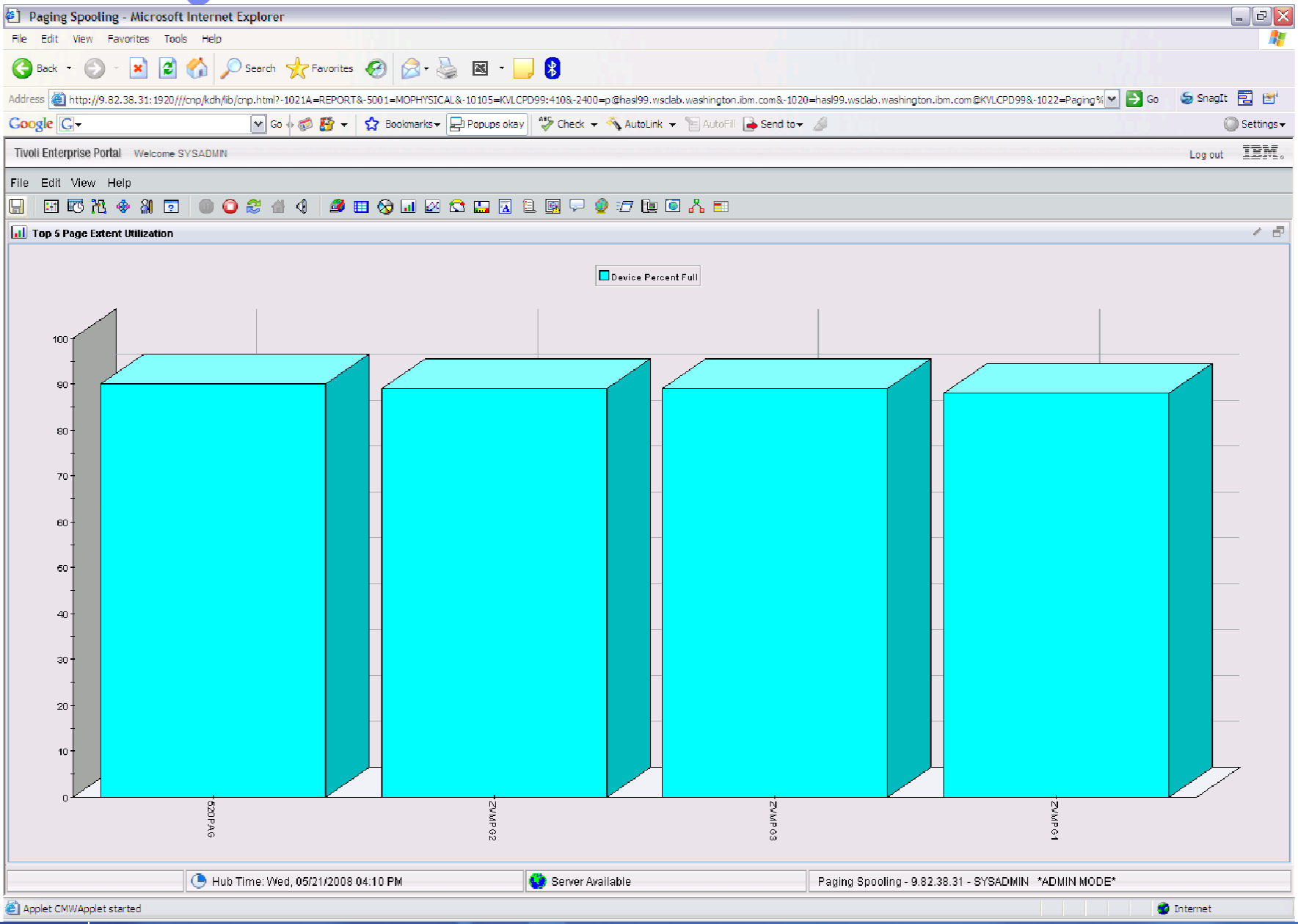
- **Plan for DASD page space utilization < 50% for performance reasons:**
  - Page space tends to get fragmented over time.
  - Large contiguous free space allows for greater paging efficiency.
  - Monitor usage with OMEGAMON XE or Q ALLOC PAGE command.
  - Block page size is the key performance indicator:
    - Aim for double digits – 10 or more pages per block set.
    - Performance Toolkit report DEV CPOWN (FCX109) "Block Page Size" Use multiple channels to spread out I/O to paging devices.
- **When Paging and Spool space fills, z/VM abends.**
  - VM warning message occurs around 90%
  - By the it's typically too late
- **Do not mix page space with any other space on a volume.**
- **Recommend using devices of the same size/geometry and performance characteristics.**
- **Calculation guidelines are located in the CP Planning and Administration Manual.**

OMEGAMON CP Owned Devices – Paging Subsystem

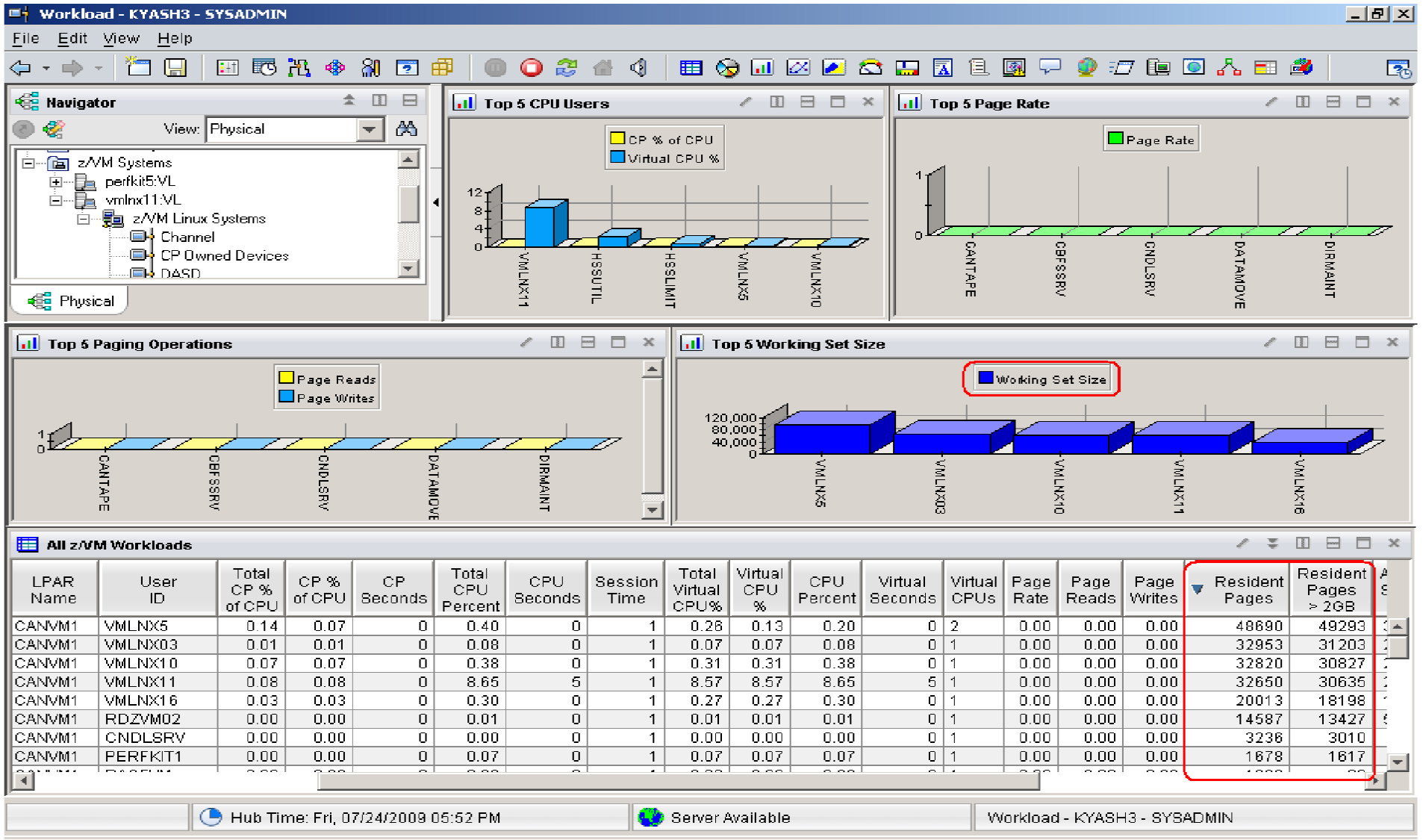


Time	System ID	LPAR Name	Device VOLSER	Device Address	PAGING SPOOLING	Allocation	Available Slots	Device Type	Device End Extent	Device Percent Full	Device Start Extent	Device Slots Used
7/14/09 18:28:17	GDLVICOM	VIC	VICRES	1C0A	DIRECT	17	2880	3390	17	6	1	180
7/14/09 18:28:17	GDLVICOM	VIC	VI1D80	1D80	PAGING	10016	1802880	3390	10016	0	1	0
7/14/09 18:28:17	GDLVICOM	VIC	VI1D81	1D81	SPOOLING	10016	1171980	3390	10016	35	1	630900
7/14/09 18:28:17	GDLVICOM	VIC	VI1D82	1D82	T-DISK	10016	1802880	3390	10016	0	1	0

z/VM Page Attributes

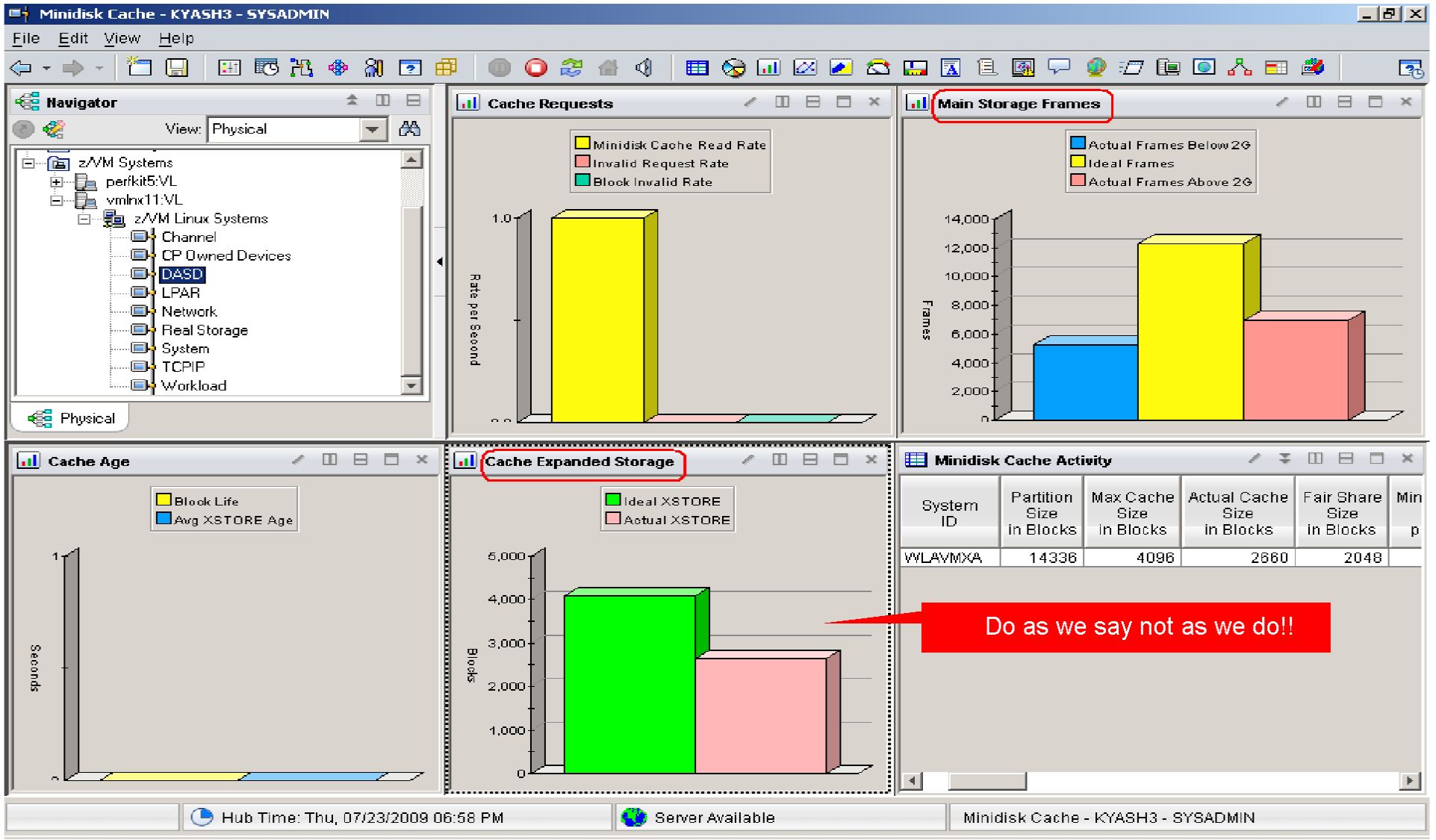


Paging – Workload Workspace

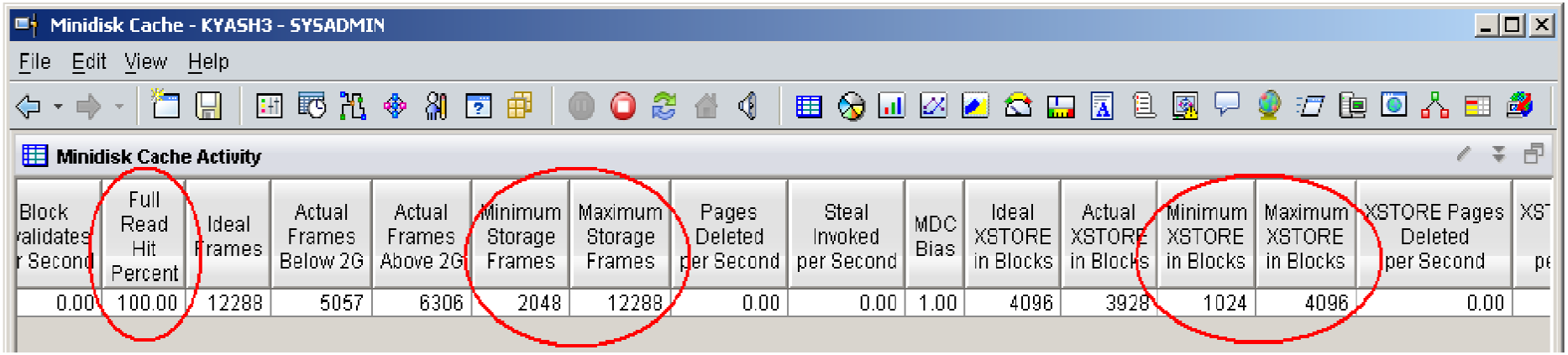


LPAR Name	User ID	Total CP % of CPU	CP % of CPU	CP Seconds	Total CPU Percent	CPU Seconds	Session Time	Total Virtual CPU%	Virtual CPU %	CPU Percent	Virtual Seconds	Virtual CPUs	Page Rate	Page Reads	Page Writes	Resident Pages	Resident Pages > 2GB
CANVM1	VMLNX5	0.14	0.07	0	0.40	0	1	0.26	0.13	0.20	0	2	0.00	0.00	0.00	48690	49293
CANVM1	VMLNX03	0.01	0.01	0	0.08	0	1	0.07	0.07	0.08	0	1	0.00	0.00	0.00	32953	31203
CANVM1	VMLNX10	0.07	0.07	0	0.38	0	1	0.31	0.31	0.38	0	1	0.00	0.00	0.00	32820	30827
CANVM1	VMLNX11	0.08	0.08	0	8.65	5	1	8.57	8.57	8.65	5	1	0.00	0.00	0.00	32650	30635
CANVM1	VMLNX16	0.03	0.03	0	0.30	0	1	0.27	0.27	0.30	0	1	0.00	0.00	0.00	20013	18198
CANVM1	RDZVM02	0.00	0.00	0	0.01	0	1	0.01	0.01	0.01	0	1	0.00	0.00	0.00	14587	13427
CANVM1	CNDLSRV	0.00	0.00	0	0.00	0	1	0.00	0.00	0.00	0	1	0.00	0.00	0.00	3236	3010
CANVM1	PERFKIT1	0.00	0.00	0	0.07	0	1	0.07	0.07	0.07	0	1	0.00	0.00	0.00	1678	1617

OMEGAMON MDISK Cache Allocations



OMEGAMON MDISK Cache Allocations – p. 2



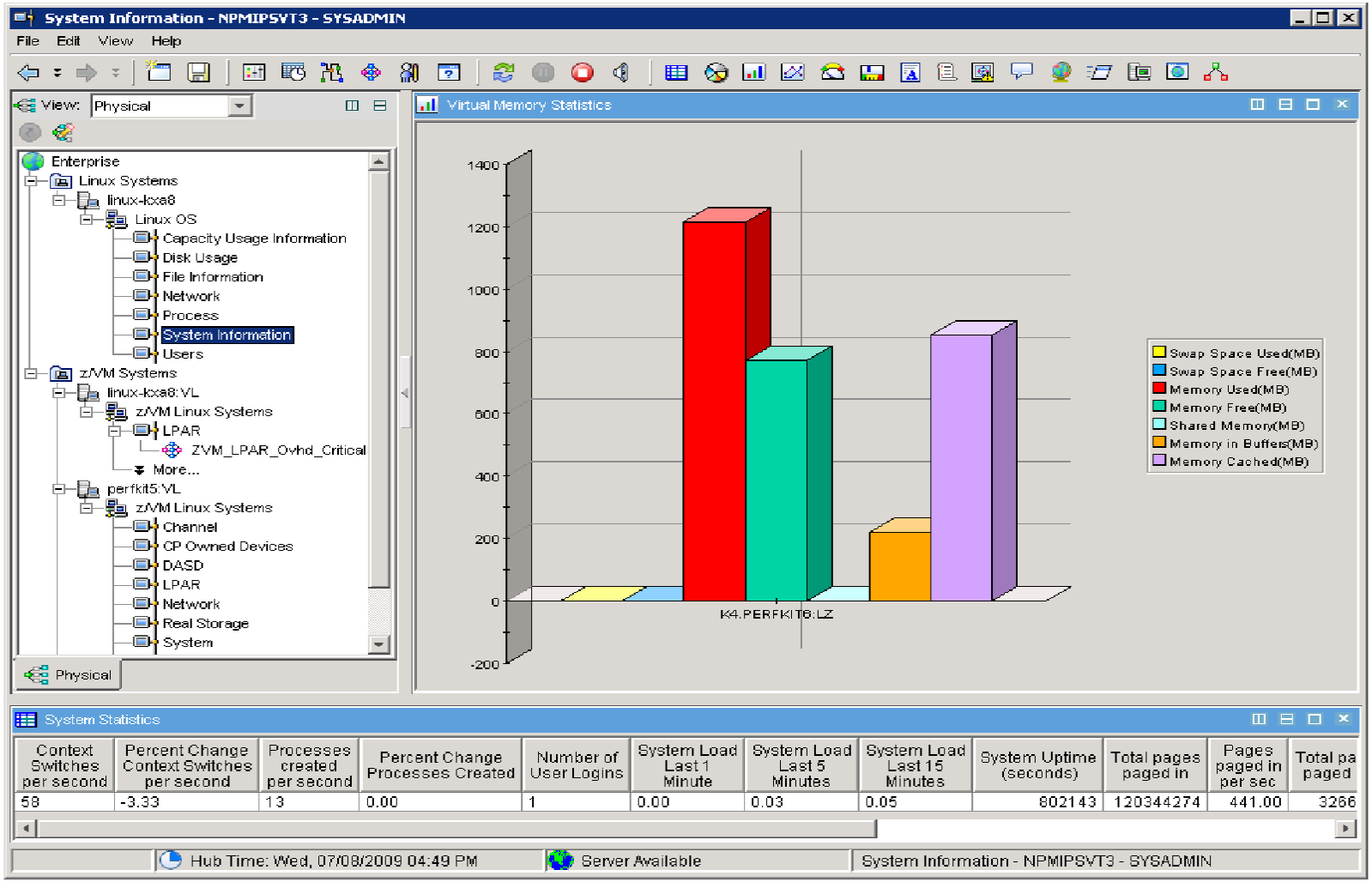
Block ...alidates r Second	Full Read Hit Percent	Ideal Frames	Actual Frames Below 2G	Actual Frames Above 2G	Minimum Storage Frames	Maximum Storage Frames	Pages Deleted per Second	Steal Invoked per Second	MDC Bias	Ideal XSTORE in Blocks	Actual XSTORE in Blocks	Minimum XSTORE in Blocks	Maximum XSTORE in Blocks	XSTORE Pages Deleted per Second	XST... pe
0.00	100.00	12288	5057	6306	2048	12288	0.00	0.00	1.00	4096	3928	1024	4096	0.00	

Is my Linux guest sized correctly?

- **More memory is not always better**
- **Excessive virtual machine sizes negatively impact performance.**
- **Linux will use all available memory**
  - Any space it doesn't need will be used as file buffer cache. Notice the large amount of cache used in example—indicates that guest may be sized too large
  - Larger Linux guests means that z/VM has to page out larger virtual machines when running other guests
- **One method—use monitor to watch for swapping. Shrink guest size until it starts swapping.**
- **Another method. Look at the Working Set Size for the Virtual Machine. This shows what z/VM is using for the guest.**
- **To handle some swapping, define a VDISK. This is much faster than swapping to a real minidisk**

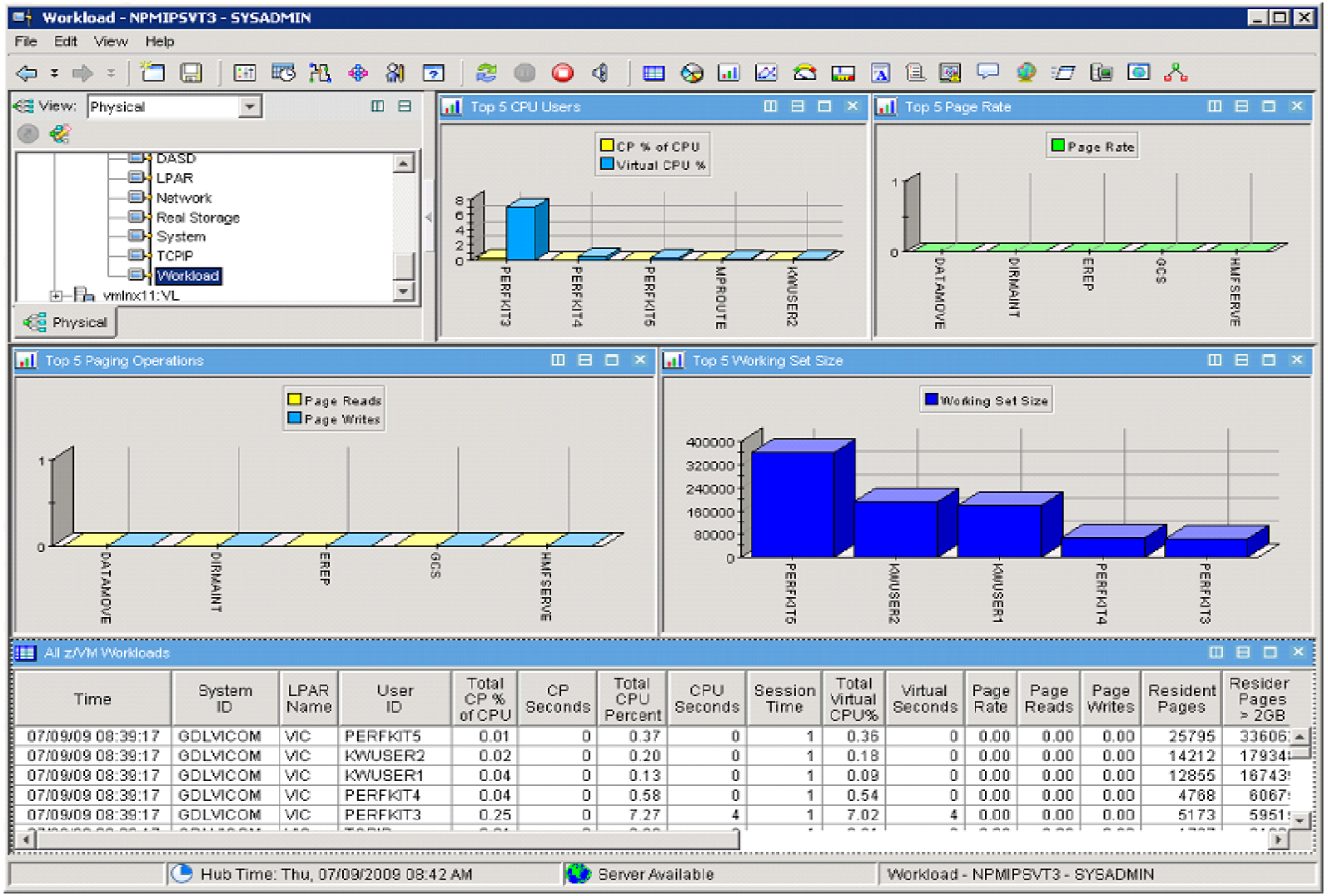
Sizing Linux Guests

Memory usage of a particular Linux virtual machine



Sizing Linux Guests

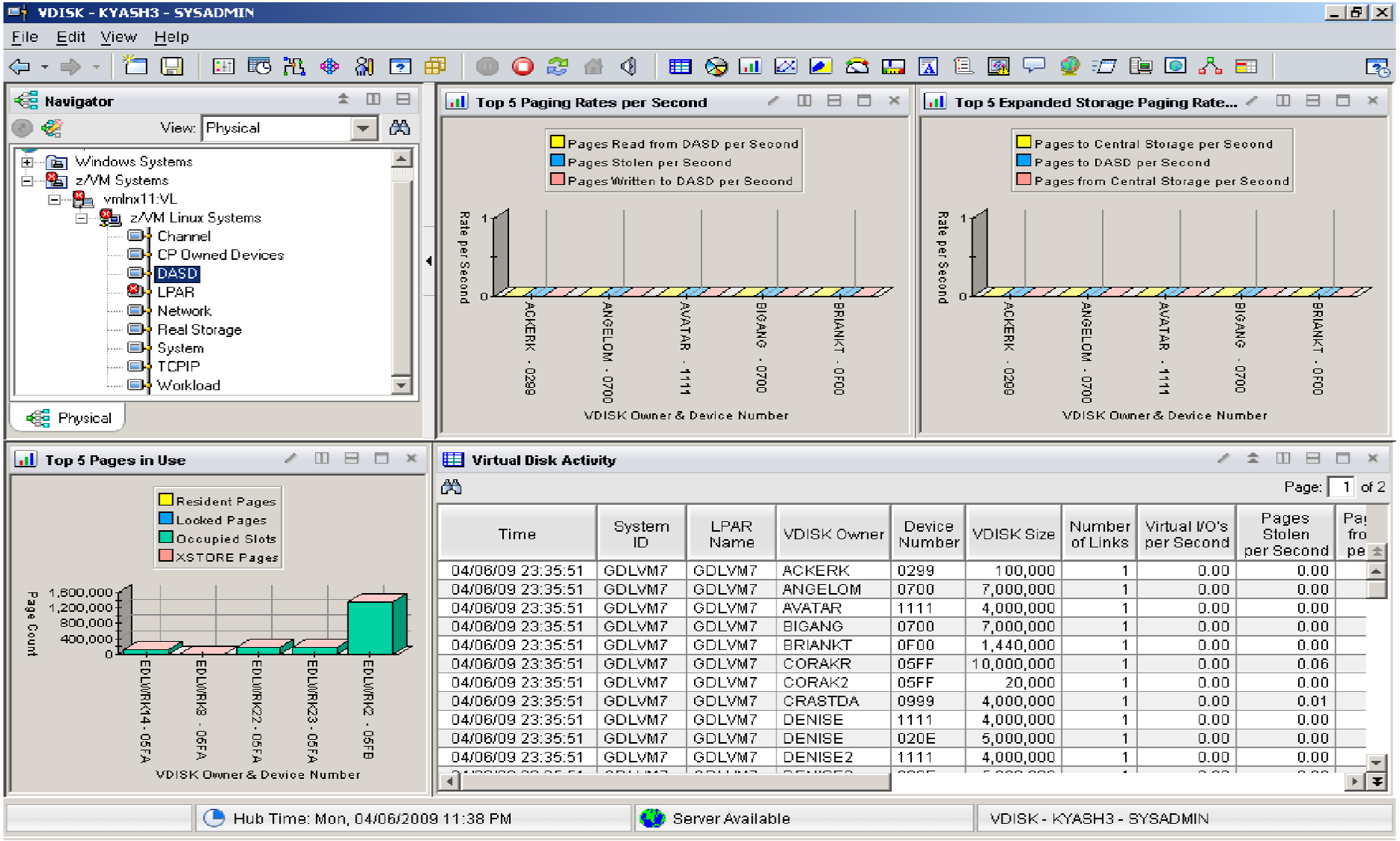
Working Set Size can be found on the Workload workspace of the z/VM agent



VDISK

- **What is it?**
  - FBA (Fixed Block Architecture disk) device emulated in-memory
    - Translation: Very fast “device”.
  - High performance paging device for Linux on z.
  - Memory is allocated by CP from the Dynamic Paging Area
  - Allocated only when referenced
    - Allocating a 10 MB device does NOT instantly consume 10 MB of pages.
    - Pages are allocated when needed.
  - Need to factor VDISK in the overall memory planning for system.

VDISK Workspace



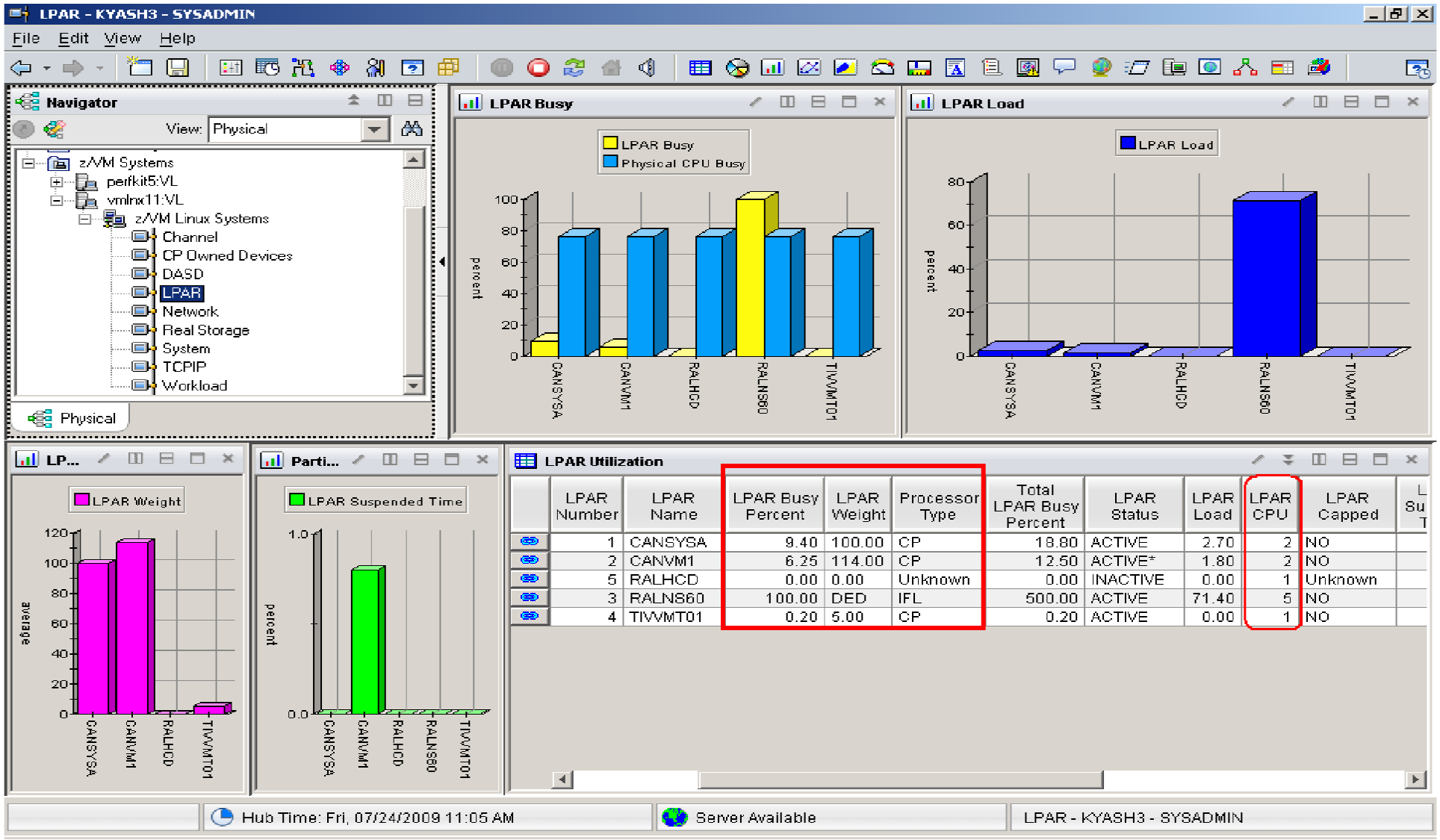
Do not ignore the hardware!

- **Just because Linux resources are virtual, do not ignore the hardware!**
  - Hardware is another potential layer of shared resources.
  - LPAR weight, CPU sharing, LPAR load, and other attributes need to be monitored for overall system performance.
  - The measurement should include the entire CEC and not just the LPAR hosting z/VM.

Processors

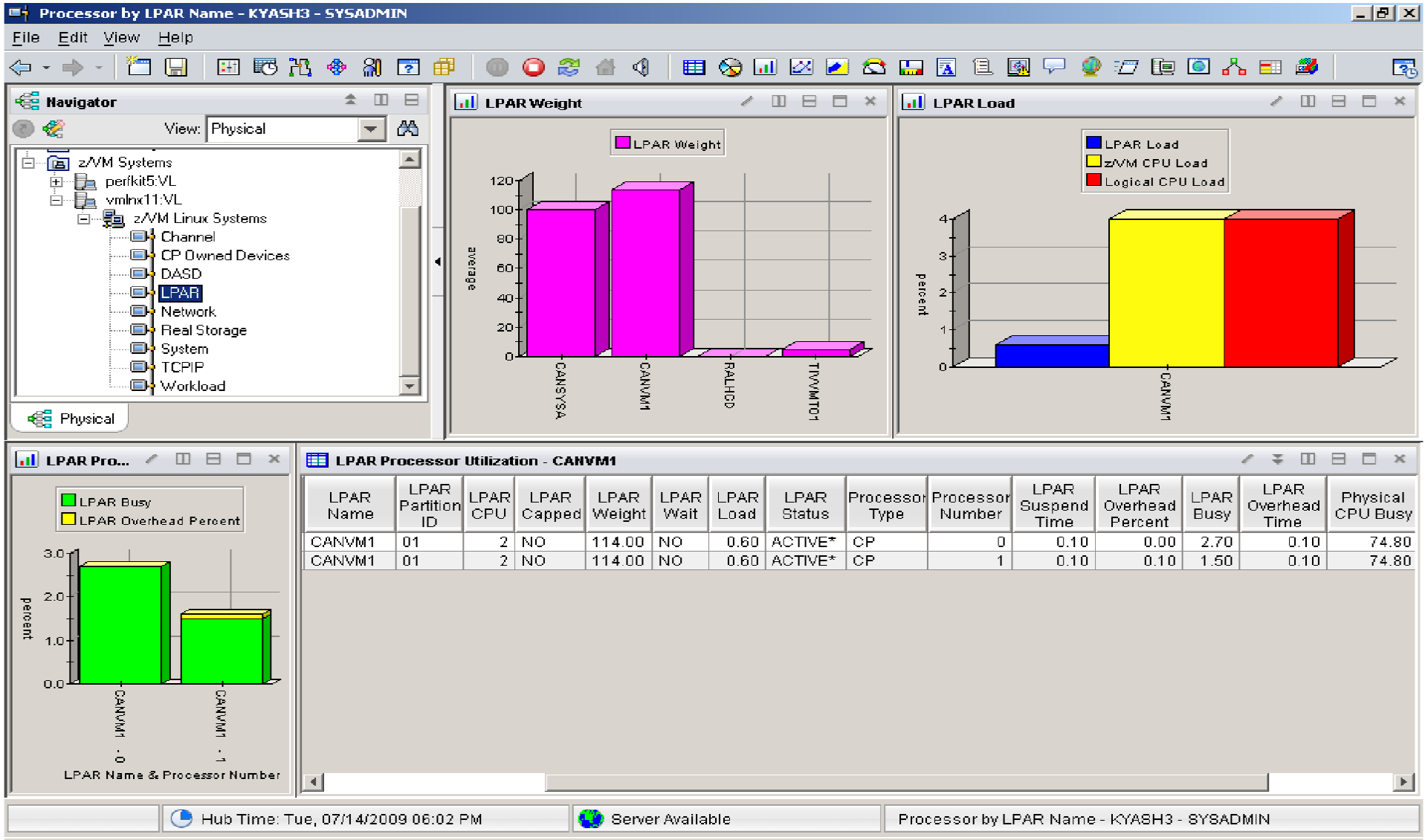
- **Logical Processors**
  - LPAR recommendation – no greater than a 4:1 logical to real ratio.
  - z/VM 5.1 and z/VM 5.2 support up to 24 processors.
  - z/VM 5.3 and z/VM 5.4 support up to 32 processors.
- **Virtual Processors**
  - Various guest systems and workloads scale differently.
  - No rule-of-thumb.
  - Virtual Machine recommendation:
    - Configure the number of virtual processors per guest for peak workload, but no more.
    - Do not define more virtual processors to a guest than logical processors defined to a z/VM LPAR.
  - High diagnose x'44' rates may be an indication of too many virtual processors.
    - Performance Toolkit reports CPU (FCX100) or PRIVOP (FCX104) can be used to monitor diagnose rates.

LPAR Utilization Workspace



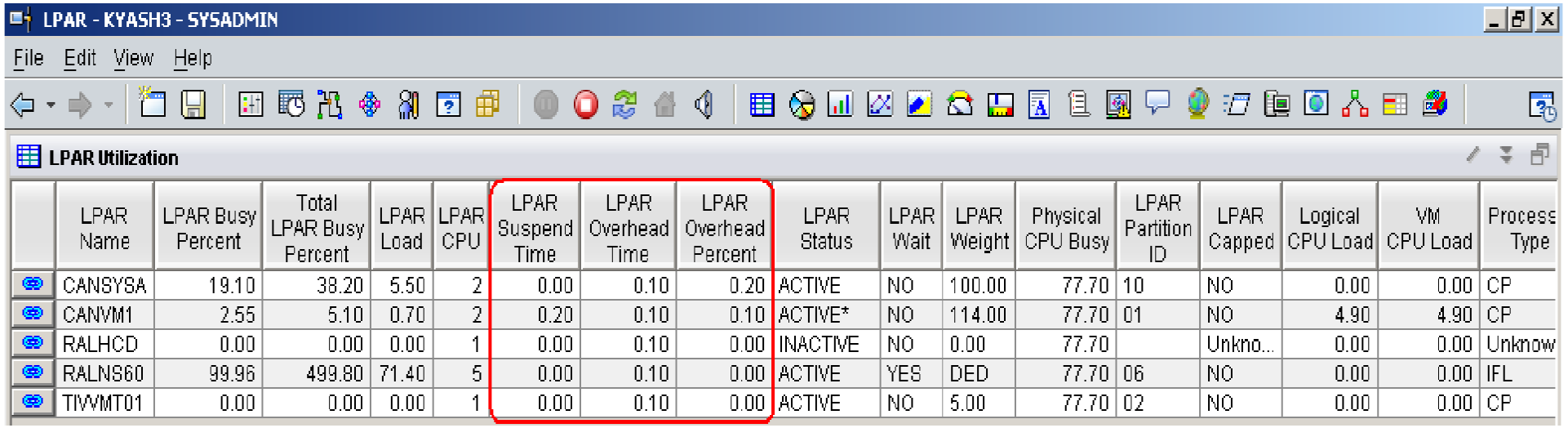
	LPAR Number	LPAR Name	LPAR Busy Percent	LPAR Weight	Processor Type	Total LPAR Busy Percent	LPAR Status	LPAR Load	LPAR CPU	LPAR Capped
	1	CANSYSA	9.40	100.00	CP	18.80	ACTIVE	2.70	2	NO
	2	CANVM1	6.25	114.00	CP	12.50	ACTIVE*	1.80	2	NO
	5	RALHCD	0.00	0.00	Unknown	0.00	INACTIVE	0.00	1	Unknown
	3	RALN860	100.00	DED	IFL	500.00	ACTIVE	71.40	5	NO
	4	TIVVMT01	0.20	5.00	CP	0.20	ACTIVE	0.00	1	NO

Processors – LPAR Processor Workspace



LPAR Name	LPAR Partition ID	LPAR CPU	LPAR Capped	LPAR Weight	LPAR Wait	LPAR Load	LPAR Status	Processor Type	Processor Number	LPAR Suspend Time	LPAR Overhead Percent	LPAR Busy	LPAR Overhead Time	Physical CPU Busy
CANVM1	01	2	NO	114.00	NO	0.60	ACTIVE*	CP	0	0.10	0.00	2.70	0.10	74.80
CANVM1	01	2	NO	114.00	NO	0.60	ACTIVE*	CP	1	0.10	0.10	1.50	0.10	74.80

LPAR Utilization Workspace – Tabular View



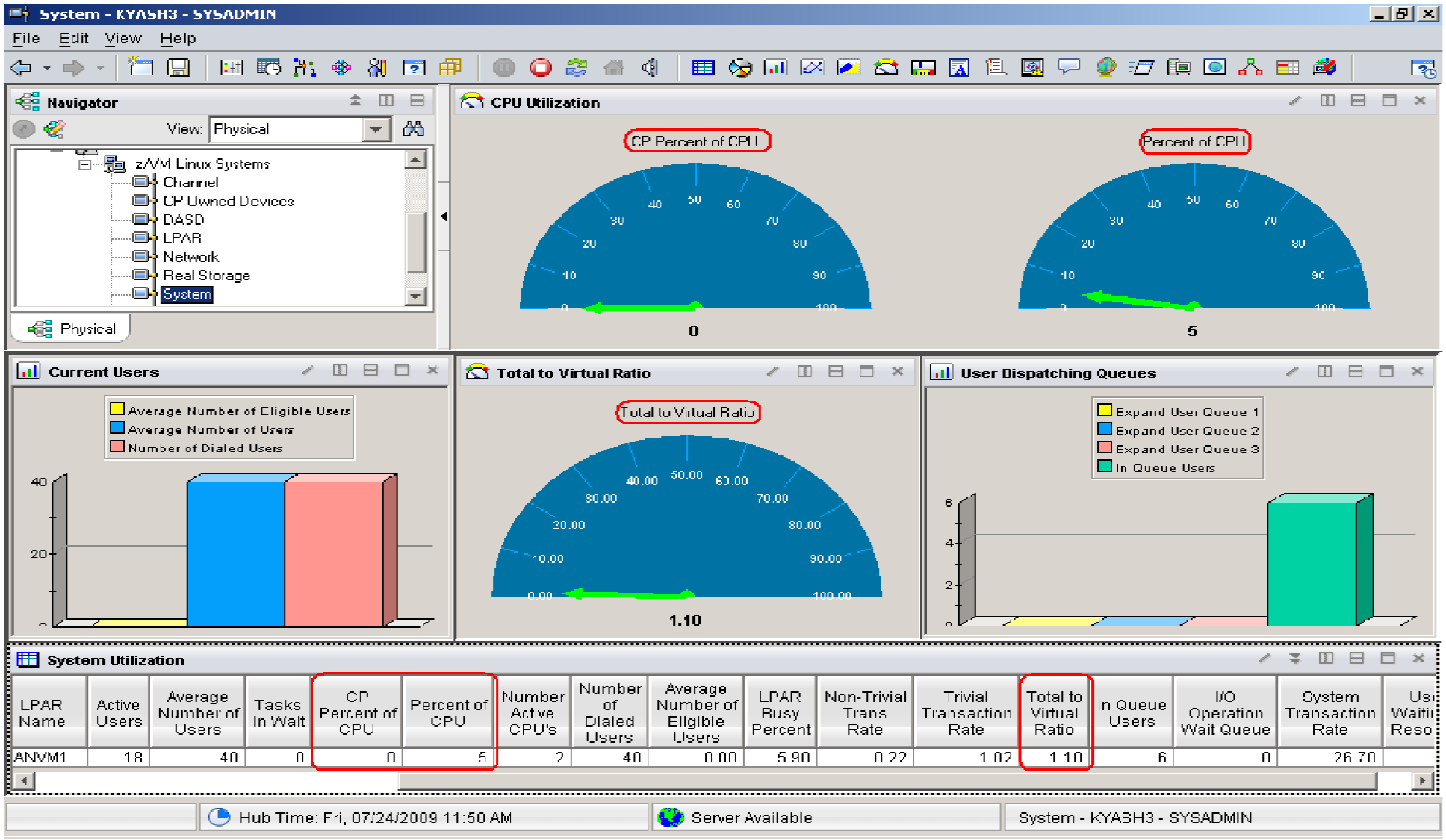
	LPAR Name	LPAR Busy Percent	Total LPAR Busy Percent	LPAR Load	LPAR CPU	LPAR Suspend Time	LPAR Overhead Time	LPAR Overhead Percent	LPAR Status	LPAR Wait	LPAR Weight	Physical CPU Busy	LPAR Partition ID	LPAR Capped	Logical CPU Load	VM CPU Load	Process Type
	CANSYSA	19.10	38.20	5.50	2	0.00	0.10	0.20	ACTIVE	NO	100.00	77.70	10	NO	0.00	0.00	CP
	CANVM1	2.55	5.10	0.70	2	0.20	0.10	0.10	ACTIVE*	NO	114.00	77.70	01	NO	4.90	4.90	CP
	RALHCD	0.00	0.00	0.00	1	0.00	0.10	0.00	INACTIVE	NO	0.00	77.70		Unkno...	0.00	0.00	Unknow
	RALNS60	99.96	499.80	71.40	5	0.00	0.10	0.00	ACTIVE	YES	DED	77.70	06	NO	0.00	0.00	IFL
	TIVVMT01	0.00	0.00	0.00	1	0.00	0.10	0.00	ACTIVE	NO	5.00	77.70	02	NO	0.00	0.00	CP

- LPAR Suspend Time: RoT: 5% Suspend time is yellow line, 10% is red line for concern.
- LPAR Overhead: This should generally be less than 5% of the Physical IFLs (CEC in an all-IFL configuration) for general LPAR management overhead, and then less than 5% of the z/VM partition IFLs.

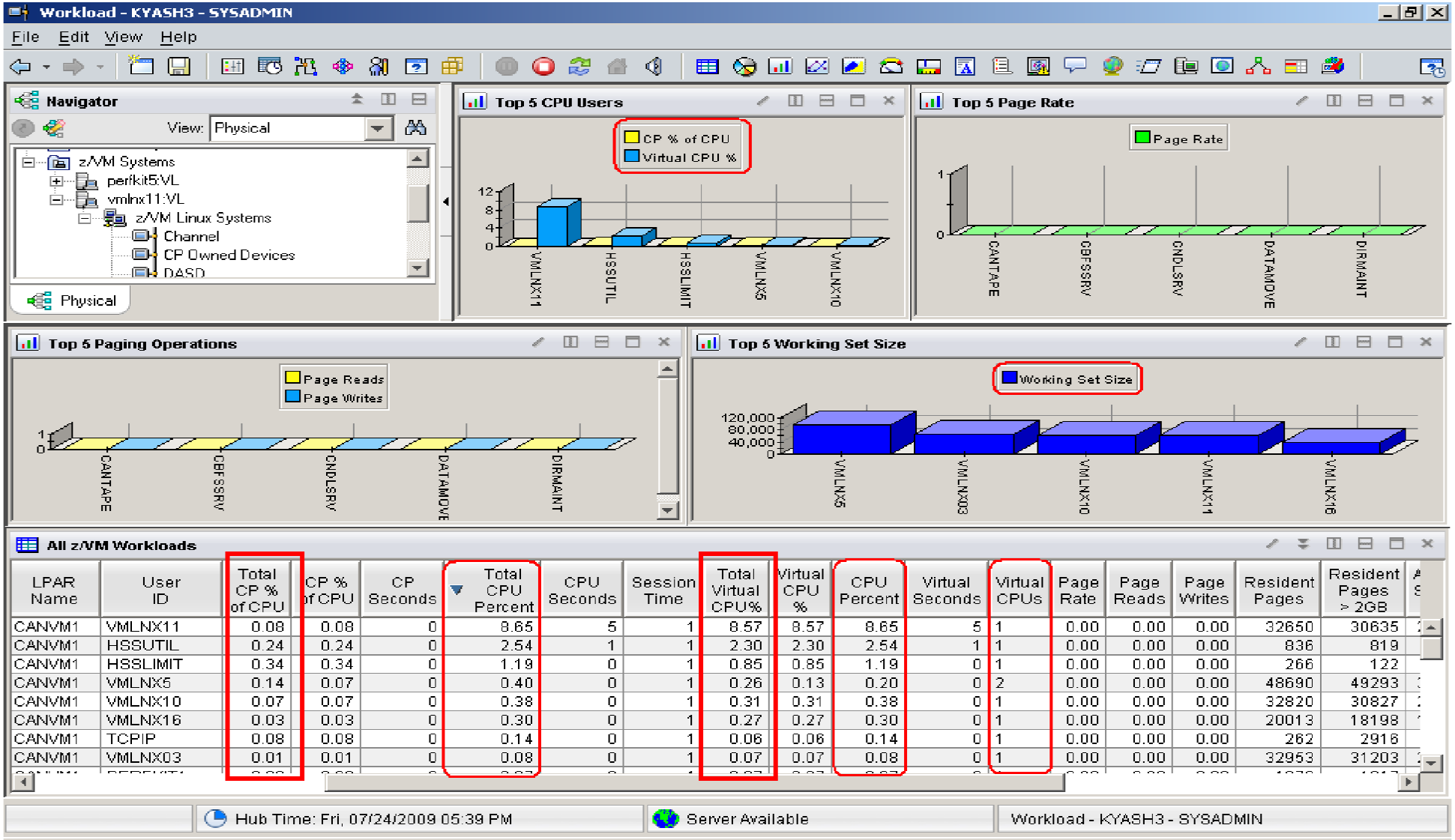
Processor Utilization

- Total Processor Utilization **This is the processor utilization from the VM perspective and includes CP, VM System, and Virtual CPU time.**
- System Time: **This is the processor time used by the VM control program for system functions that are not directly related to any one virtual machine. This should be less than 10% of the total.**
- CP Processor Time: **This is the processor time used by the VM control program in support of individual virtual machines.**
- Virtual Processor Time: (Emulation Time): **This is processor time consumed by the virtual machine and the applications within it.**
- Total to Virtual Ratio **The ratio of total processor time to virtual processor time is often used as an indicator of z/VM efficiency or overhead. The closer to 1.0, the better the z/VM efficiency. RoT: Should explore causes of a ratio over 1.30.**

System Processor Utilization Workspace



Processors – Workload Workspace

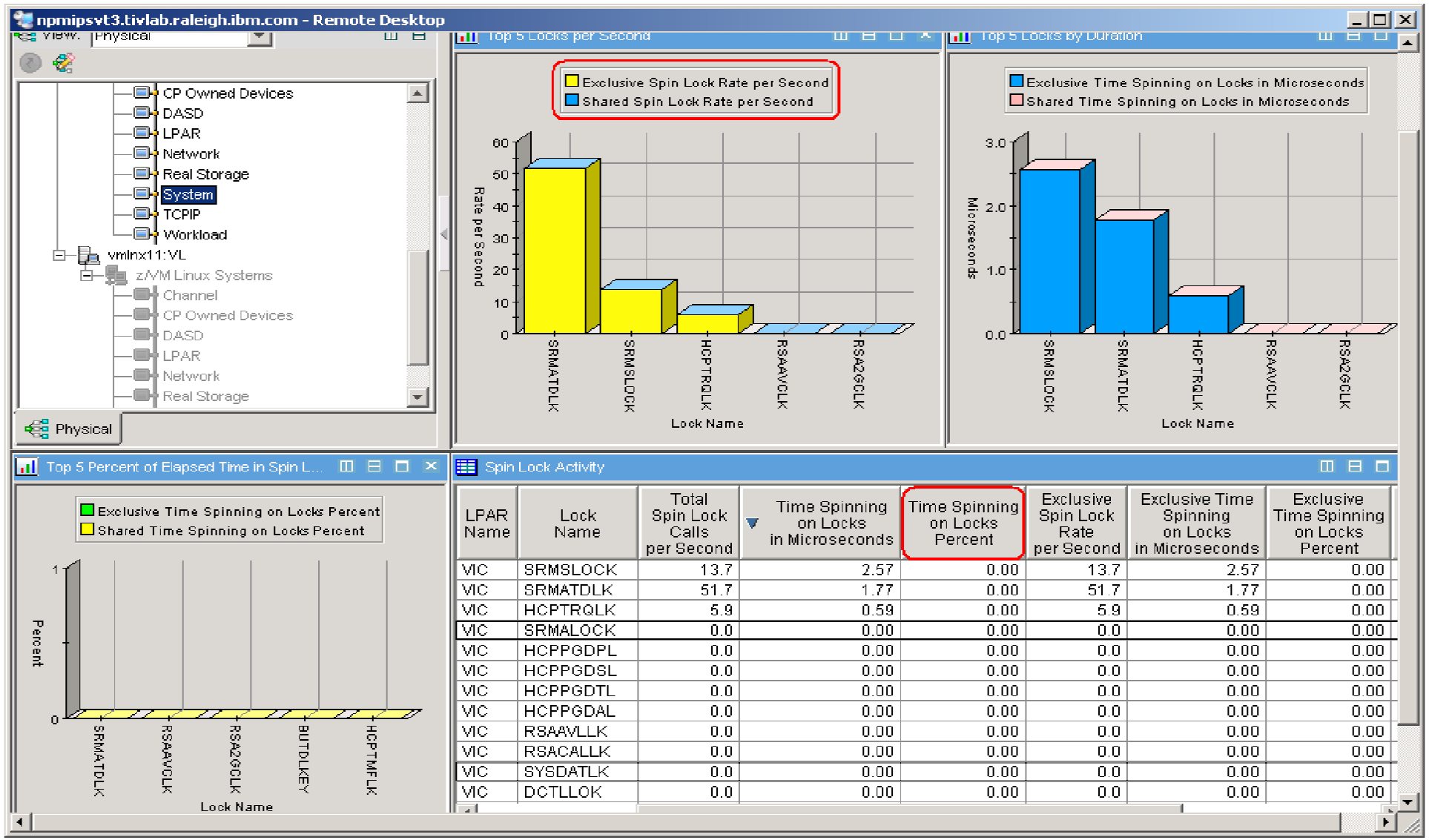


LPAR Name	User ID	Total CP % of CPU	CP % of CPU	CP Seconds	Total CPU Percent	CPU Seconds	Session Time	Total Virtual CPU%	Virtual CPU %	CPU Percent	Virtual Seconds	Virtual CPUs	Page Rate	Page Reads	Page Writes	Resident Pages	Resident Pages > 2GB
CANVM1	VMLNX11	0.08	0.08	0	8.65	5	1	8.57	8.57	8.65	5	1	0.00	0.00	0.00	32650	30635
CANVM1	HSSUTIL	0.24	0.24	0	2.54	1	1	2.30	2.30	2.54	1	1	0.00	0.00	0.00	836	819
CANVM1	HSSLIMIT	0.34	0.34	0	1.19	0	1	0.85	0.85	1.19	0	1	0.00	0.00	0.00	266	122
CANVM1	VMLNX5	0.14	0.07	0	0.40	0	1	0.26	0.13	0.20	0	2	0.00	0.00	0.00	48690	49293
CANVM1	VMLNX10	0.07	0.07	0	0.38	0	1	0.31	0.31	0.38	0	1	0.00	0.00	0.00	32820	30827
CANVM1	VMLNX16	0.03	0.03	0	0.30	0	1	0.27	0.27	0.30	0	1	0.00	0.00	0.00	20013	18198
CANVM1	TCPIP	0.08	0.08	0	0.14	0	1	0.06	0.06	0.14	0	1	0.00	0.00	0.00	262	2916
CANVM1	VMLNX03	0.01	0.01	0	0.08	0	1	0.07	0.07	0.08	0	1	0.00	0.00	0.00	32953	31203

Spin Lock Wait

- Time Spinning on Locks Percent:
  - The percentage of time processors spend spinning on formal spin locks. RoT: Should be less than 10%.
  - Increases as number of logical processors increases.

Spinlock Workspace

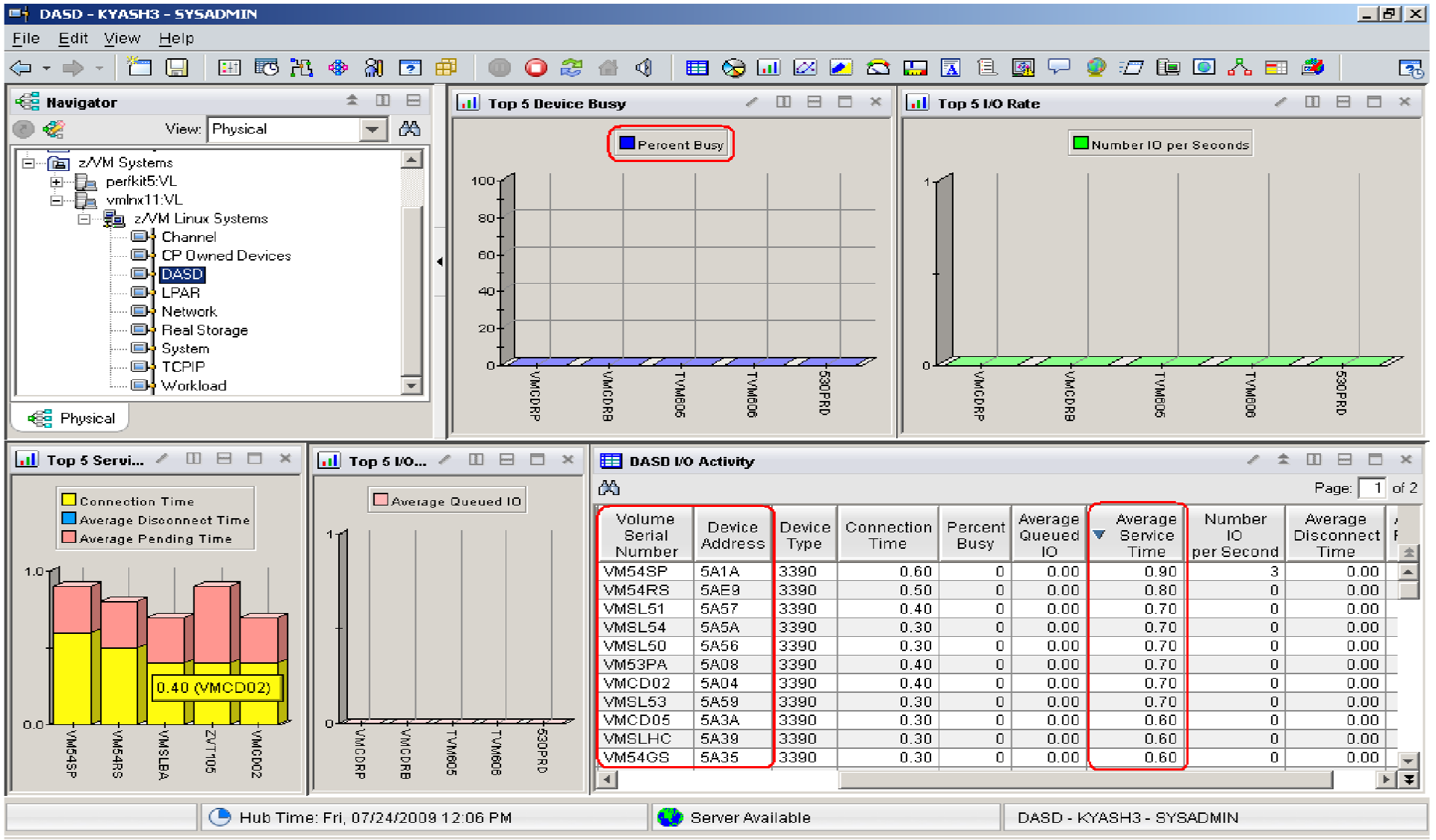


LPAR Name	Lock Name	Total Spin Lock Calls per Second	Time Spinning on Locks in Microseconds	Time Spinning on Locks Percent	Exclusive Spin Lock Rate per Second	Exclusive Time Spinning on Locks in Microseconds	Exclusive Time Spinning on Locks Percent
VIC	SRMSLOCK	13.7	2.57	0.00	13.7	2.57	0.00
VIC	SRMATDLK	51.7	1.77	0.00	51.7	1.77	0.00
VIC	HCPTRQLK	5.9	0.59	0.00	5.9	0.59	0.00
VIC	SRMALOCK	0.0	0.00	0.00	0.0	0.00	0.00
VIC	HCPPGDPL	0.0	0.00	0.00	0.0	0.00	0.00
VIC	HCPPGDSL	0.0	0.00	0.00	0.0	0.00	0.00
VIC	HCPPGDTL	0.0	0.00	0.00	0.0	0.00	0.00
VIC	HCPPGDAL	0.0	0.00	0.00	0.0	0.00	0.00
VIC	RSAAVLLK	0.0	0.00	0.00	0.0	0.00	0.00
VIC	RSACALLK	0.0	0.00	0.00	0.0	0.00	0.00
VIC	SYSDATLK	0.0	0.00	0.00	0.0	0.00	0.00
VIC	DCTLLOK	0.0	0.00	0.00	0.0	0.00	0.00

Direct Access Storage Devices (DASD)

- Avg Pending Time for DASD
  - Average pending time for real DASD I/Os. RoT: Should be less than 1 millisecond.
- **Items worth keeping an eye on:**
  - **Number of I/O's per Second, Percent Busy**
  - **Avg Service Time** Average service time for real DASD devices (sum of the pending, connect, and disconnect times).
  - **DASD I/O Rate** Rate of traditional real I/Os per second to real DASD devices. Worth monitoring.

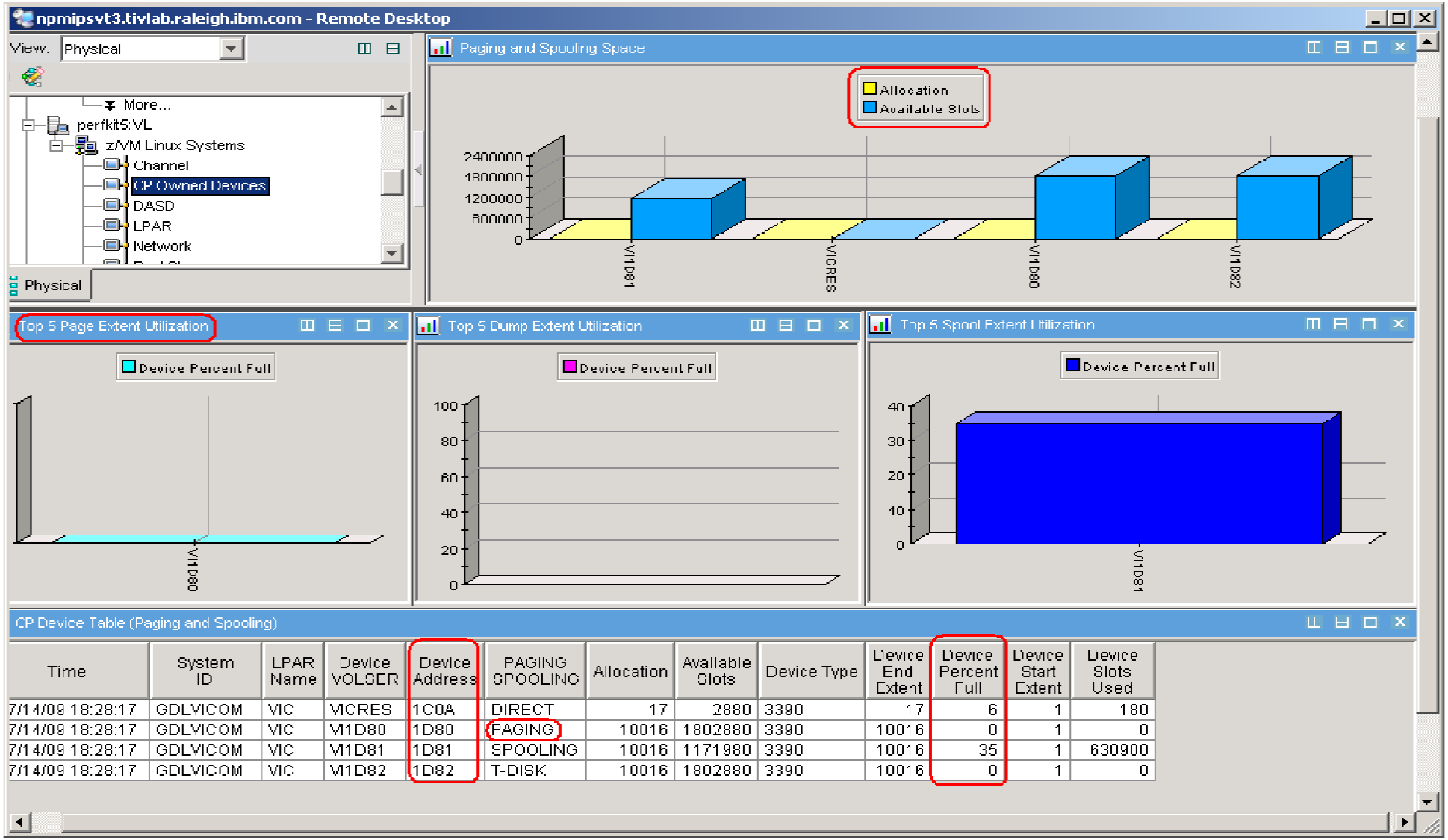
DASD I/O Workspace



System Dump & Spool Space

- **Dump Space**
  - Ensure there is sufficient dump space defined to the system.
  - Dump space requirements vary according to memory usage.
    - Q DUMP – identifies allocated dump space.
    - Calculation guidelines are located in CP Planning and Administration Manual.
- **Spool Space**
  - Various uses:
    - User printer, punch, reader files (console logs)
    - DCSS, NSS
    - System files
    - Page space overflow
  - Spool Management:
    - Monitor with Q ALLOC SPOOL command, OMEGAMON XE or Operations Manager for z/VM
    - command.
    - SFPURGER utility:
      - Rule based tool to clean up spool space.
      - Included in the no charge CMS Utilities Feature (CUF).

System Dump & Spool Space



Time	System ID	LPAR Name	Device VOLSER	Device Address	PAGING SPOOLING	Allocation	Available Slots	Device Type	Device End Extent	Device Percent Full	Device Start Extent	Device Slots Used
7/14/09 18:28:17	GDLVICOM	VIC	VICRES	1C0A	DIRECT	17	2880	3390	17	6	1	180
7/14/09 18:28:17	GDLVICOM	VIC	VI1D80	1D80	PAGING	10016	1802880	3390	10016	0	1	0
7/14/09 18:28:17	GDLVICOM	VIC	VI1D81	1D81	SPOOLING	10016	1171980	3390	10016	35	1	630900
7/14/09 18:28:17	GDLVICOM	VIC	VI1D82	1D82	T-DISK	10016	1802880	3390	10016	0	1	0

System Resource Management Settings

- Influence the z/VM scheduler and dispatcher behavior.
- Default values are an artifact from the past:
  - Interactive CMS virtual machines
  - Small memory footprints
- STORBUF
  - Defines amount of memory to be used in scheduler algorithms.
  - Recommend modification to over-commit central storage.
    - Default values - STORBUF 120 105 95
    - Recommended starting values - STORBUF 300 250 200
- LDUBUF
  - Defines amount of paging “capacity” to be used in scheduler algorithms.
  - There are conflicting opinions on a recommended setting:
    - Default values - LDUBUF 100 75 60
    - Default values may be “OK” as a starting point depending on:
      - Amount of DASD paging capacity and number of volumes defined
      - Number and size of active Linux guests
- DSPBUF
  - Defines number of guests allowed in the dispatch list:
    - Default values - DSPBUF 32767 32767 32767
    - Not recommended to adjust these settings unless directed by development

Quick Dispatch

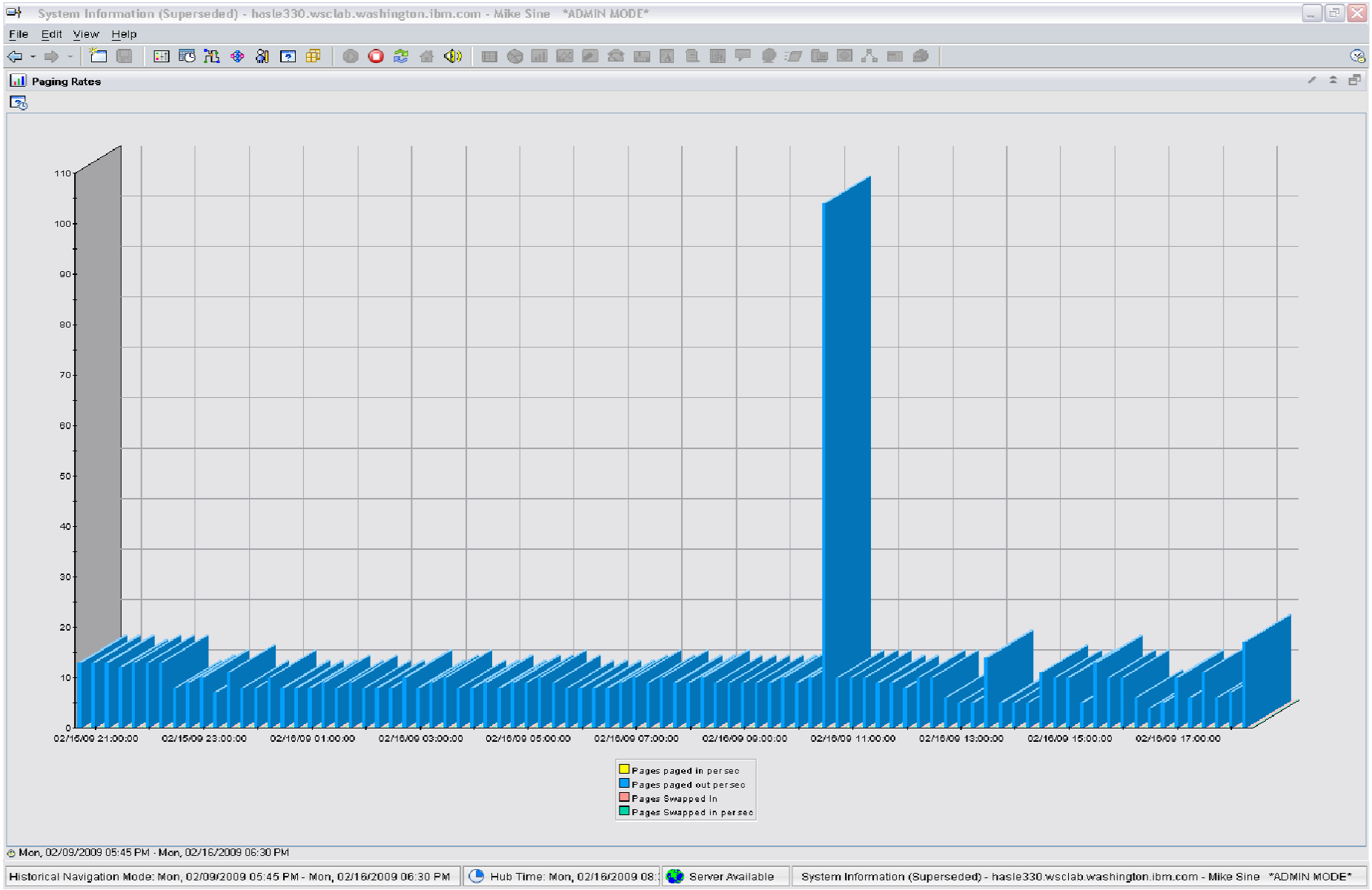
- **Setting QUICKDSP:**
  - Bypasses System Resource Management controls.
  - Places a virtual machine directly into the dispatch list.
  - Virtual Machine exempt from being held back in an eligible list.
- **QUICKDSP should be reserved for:**
  - Select production guests only.
  - Service Virtual Machines performing critical functions on behalf of other guests (i.e. RACF, TCPIP).
- **SRM values should be used to adjust scheduler/dispatcher behavior for servicing most guests.**
- **See** http://www2.marist.edu/htbin/wlvtype?LINUX-VM.30359 **for an excellent detailed explanation by Malcolm Beattie (IBM).**

Persistent Historical Views

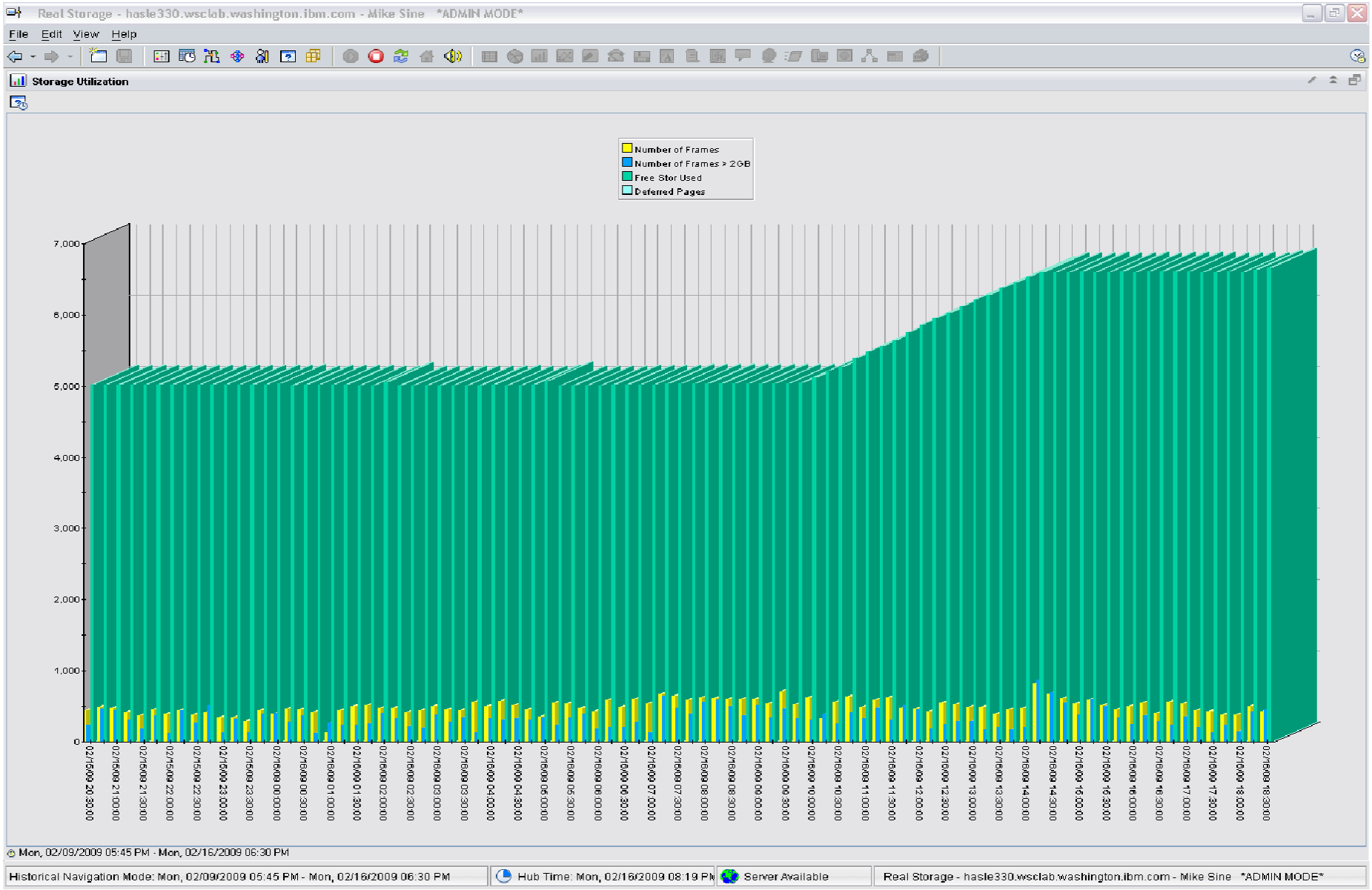
This makes it easier to see anomalies, or match spikes. Capturing performance data as a base line is a must:

- **General history data – business as usual.**
- **Detailed raw monitor data prior to and following any major changes.**
- **Ability to review attributes of a past incident.**

Persistent Historical Views



Persistent Historical Views



New Tivoli Common Reporting (TCR)

- **TCR reports available on the OPAL website**
  - http://www-18.lotus.com/wps/portal/topal
- **What is TCR?**
  - Tivoli Common Reporting.
  - Consistent approach to viewing and administering reports.
  - Built on top of open source reporting tool called: BIRT.
  - Flexible development environment (Eclipse based) for creating report definitions.
  - Five templates provided for download.
  - Taking suggestions for more

Sample Reports Available

- **z/VM VM System CPU Utilization**
- **z/VM VM System Paging Utilization**
- **z/VM Linux System CPU Utilization**
- **z/VM VM System CP-Owned Device Utilization**
- **z/VM VM System TCP Server Statistics**

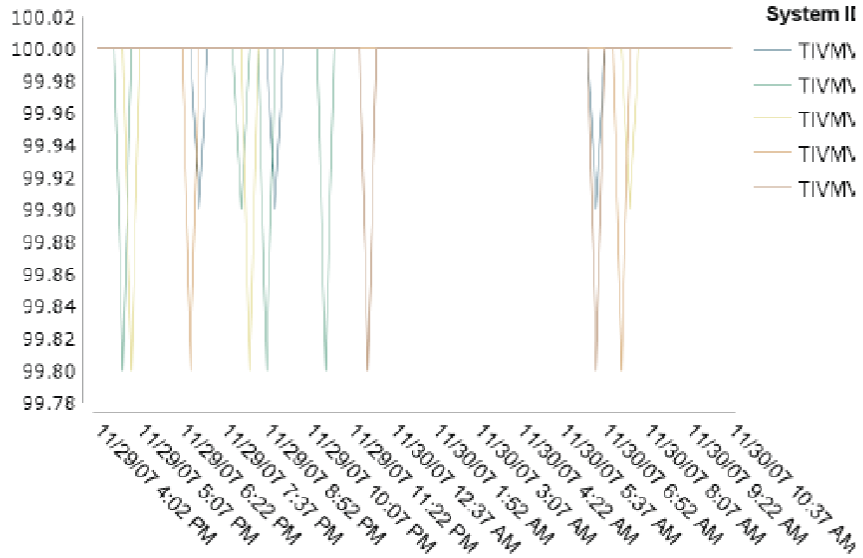
Tivoli

IBM

z/VM System CPU Utilization

Report Period	**All**	**Significant Resources Selected**	**5**
Start Date	**Dec 31, 1969 12:00 AM**	**End Date**	**Nov 30, 2007 11:59 PM**
System ID	**All**	**LPAR Name**	**All**

LPAR Busy



Available Summarization Time Periods:
Hourly
Daily
Weekly
Monthly
Not Summarized Data

System = TIVMVS6

LPAR Name	LPAR Busy	LPAR Load	LPAR Suspend Time	LPAR Overhead Time	Date/Time

November 30, 2007 2:26:24 PM EST 1 / 18

Tivoli

IBM

System = TIVMVS6

LPAR Name	LPAR Busy	LPAR Load	LPAR Suspend Time	LPAR Overhead Time	Date/Time
RALNS31	100	4.2	0	.6	Nov 29, 2007 4:02 PM
RALNS32	100	4.2	0	.6	Nov 29, 2007 4:02 PM
RALNS61	100	4.2	0	.6	Nov 29, 2007 4:02 PM
TIVMVS1	100	2.09	0	.6	Nov 29, 2007 4:02 PM
TIVMVS10	100	2.09	0	.6	Nov 29, 2007 4:02 PM
RALNS31	100	4.2	0	.6	Nov 29, 2007 4:08 PM
RALNS32	100	4.2	0	.6	Nov 29, 2007 4:08 PM
RALNS61	100	4.2	0	.6	Nov 29, 2007 4:08 PM
TIVMVS1	100	2.09	0	.6	Nov 29, 2007 4:08 PM
TIVMVS10	100	2.09	0	.6	Nov 29, 2007 4:08 PM
RALNS31	100	4.2	0	.6	Nov 29, 2007 4:22 PM
RALNS32	100	4.2	0	.6	Nov 29, 2007 4:22 PM
RALNS61	100	4.2	0	.6	Nov 29, 2007 4:22 PM
TIVMVS1	100	2.09	0	.6	Nov 29, 2007 4:22 PM
TIVMVS10	100	2.09	0	.6	Nov 29, 2007 4:22 PM
RALNS31	100	4.2	0	.6	Nov 29, 2007 4:37 PM
RALNS61	100	4.2	0	.6	Nov 29, 2007 4:37 PM
TIVMVS1	100	2.09	0	.6	Nov 29, 2007 4:37 PM
TIVMVS10	100	2.09	0	.6	Nov 29, 2007 4:37 PM
RALNS32	99.8	4.2	0	.6	Nov 29, 2007 4:37 PM
RALNS31	100	4.2	0	.6	Nov 29, 2007 4:52 PM
RALNS32	100	4.2	0	.6	Nov 29, 2007 4:52 PM
TIVMVS1	100	2.09	0	.6	Nov 29, 2007 4:52 PM
TIVMVS10	100	2.09	0	.6	Nov 29, 2007 4:52

November 30, 2007 2:26:51 PM EST 2 / 18

OMEGAMON XE on z/VM and Linux
An Integrated Monitoring Approach

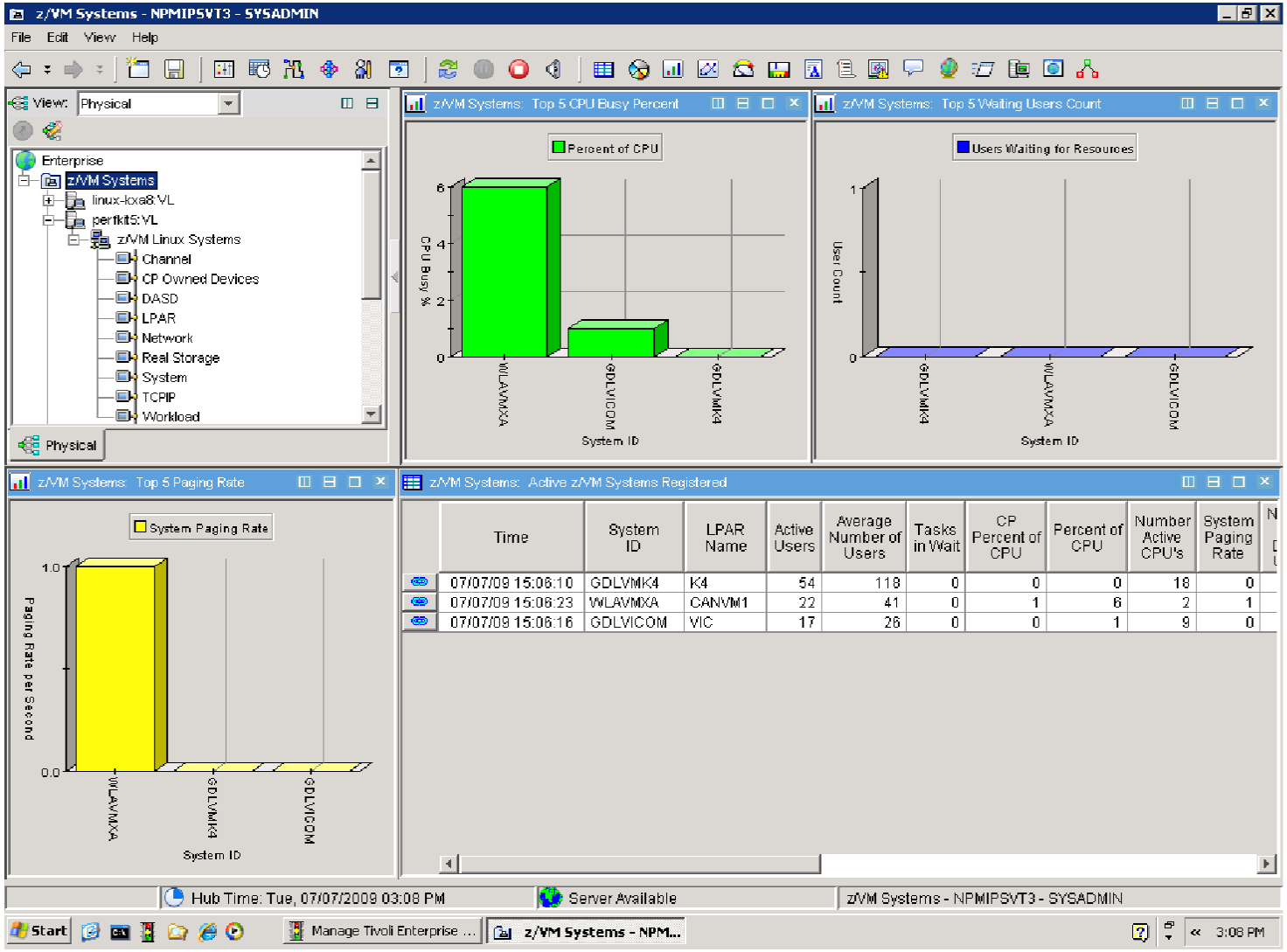
- **Provides performance monitoring for z/VM and Linux guests**
- **Linux agents gather performance data from Linux guests**
- **z/VM agent gathers performance data from z/VM**
  - Including z/VM view of guests
  - Uses IBM Performance Toolkit for VM as its data source
  - Linux provides APPLDATA to CP monitor
- **Executes automated actions in response to defined events or situations**
- **Part of the Tivoli Management Services infrastructure and OMEGAMON family of products**
  - Specifically focused on z/VM and Linux guests

Use Cases

- **Overall health of your z/VM systems**
- **Managing CP Owned Volumes**
- **CPU usage**
- **System running slowly—how to debug**
- **Sizing Linux guests**

Scenario—Overall Health of Your System

At a quick glance you can see the %CPU usage, number of users in a wait state, and paging rates of all your z/VM systems



Scenario—Overall Health of Your System

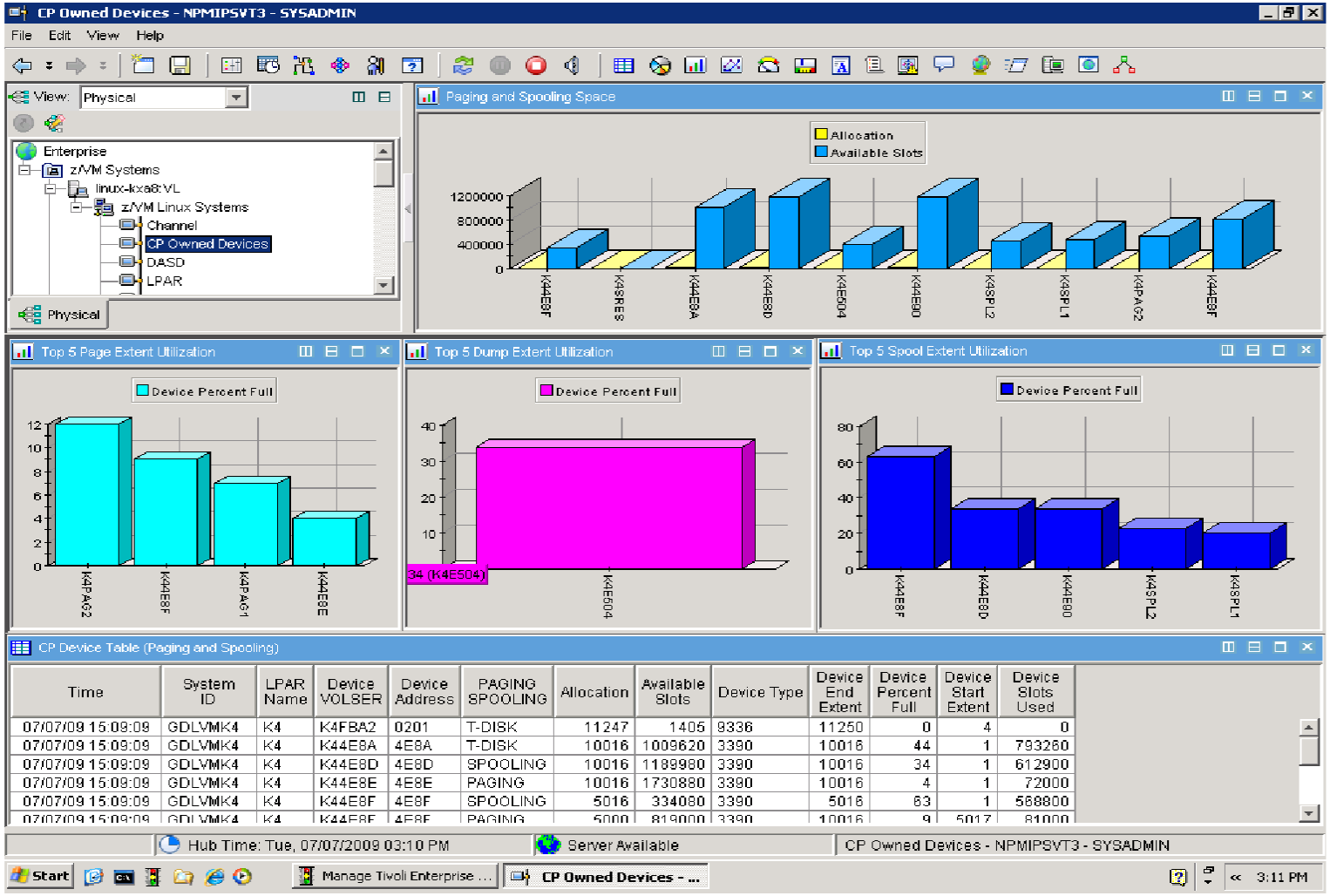
- **Things to look for**
  - CPU usage
    - Is any one system using more CPU than expected
    - Is any one system using less CPU than expected—you may have an underutilized processor and be wasting capacity
    - Remember, a DEDICATED processor will show 100%
  - Users waiting for resources
    - Number of users at the end of the monitoring interval who are either in:
      - Eligible list—waiting to enter the dispatch list
      - Nondispatchable
        - Waiting for paging
        - Waiting for I/O completion
      - Dispatchable
        - Waiting for a processor

Scenario—Overall Health of Your System

- **Things to look for**
  - System paging rate
    - Number of page reads per second
    - Not a complete indicator of your paging effectiveness, but a good first glance
      - If the rate is low, and you don't have many users waiting or paging to complete (dispatch list), then you don't have a problem
      - If rate is low and you DO have many users in dispatch list, it may be an indication of a paging problem.
        - > High dispatch list number could be for other reasons such as I/O contention. You need to check.
    - If the rate is high, then you may need to tune your paging subsystem.

Scenario—CP Owned Volumes

One place to check to see if you need to add more space on your CP-Owned volumes



Scenario—CP Owned Volumes

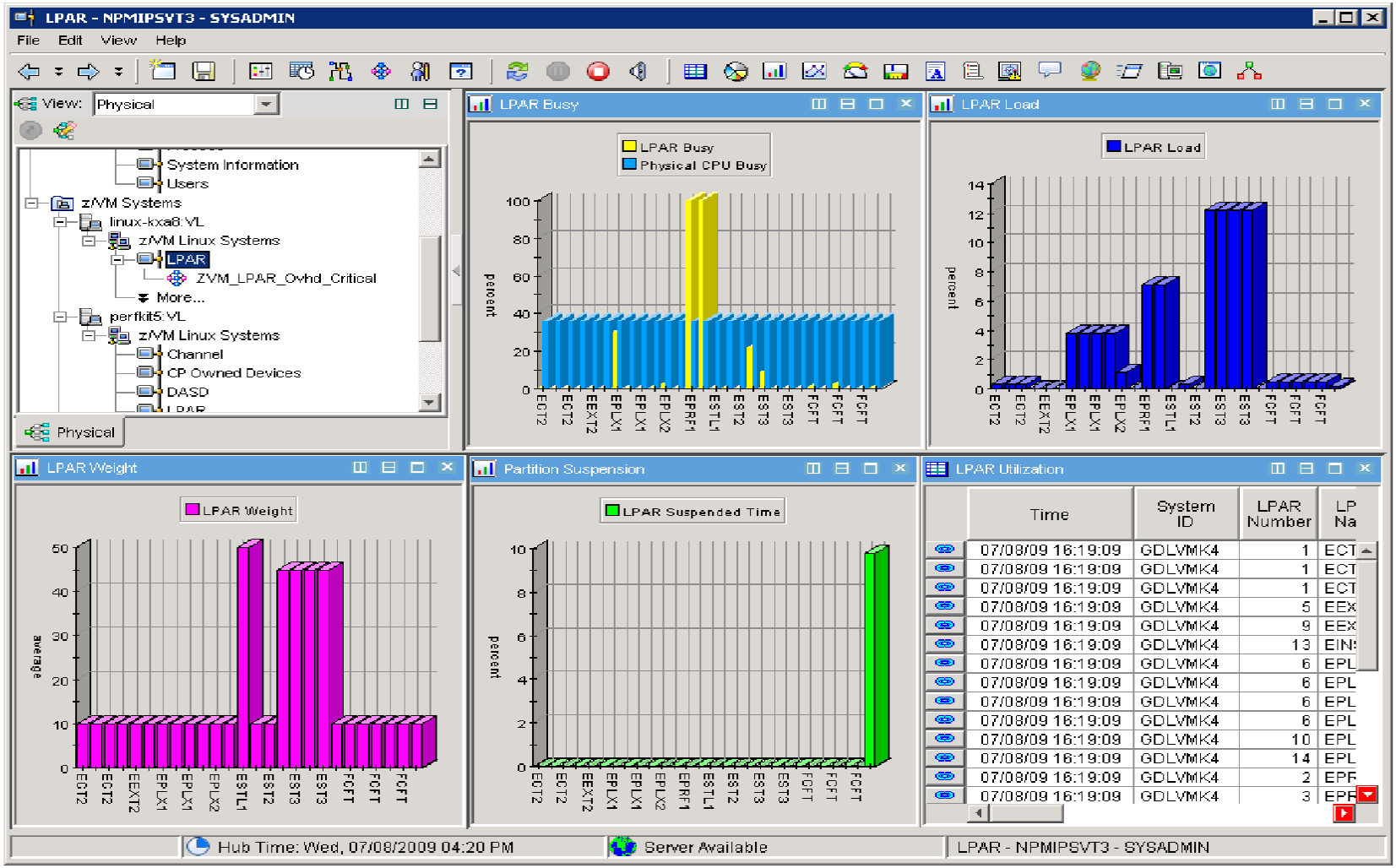
- **CP-Owned Volumes**
  - Consist of
    - Page, Spool, Tdisk, Directory and Dump (subset of spool)
  - Considerations
    - Page
      - Are my volumes getting full—do I need to add more
      - Is my paging spread out sufficiently
    - Spool
      - Used mainly for Reader and print files, Dumps and Named Saved Systems
      - If full, can I delete old spool files, or do I need more
      - Good to have an automatic cleanup program, perhaps based on age o spool file (SFPURGER utility)
    - Tdisk
      - Add more volumes, or free up existing space

Scenario—CP Owned Volumes

- **General tips**
  - Page space utilization should always be < 50%
  - Never put Paging and Spool space on the same volume
  - Allocate Spool and Page volumes to try and reduce I/O contention by separating them as much as possible (control unit, channel, etc)
  - Dedicated paging devices reduce contention for paging
  - Try to avoid putting highly used files on the same volume as paging and spool space, such as the CMS system disk
  - Use your fastest devices for Paging
  - Multiple Paging devices allow more overlap of paging operations
  - Expanded storage can be used for paging
  - Directory space is not heavily used, can be placed anywhere

Scenario—CPU Usage

LPAR workspace allows you to look at all your LPARs across the CEC

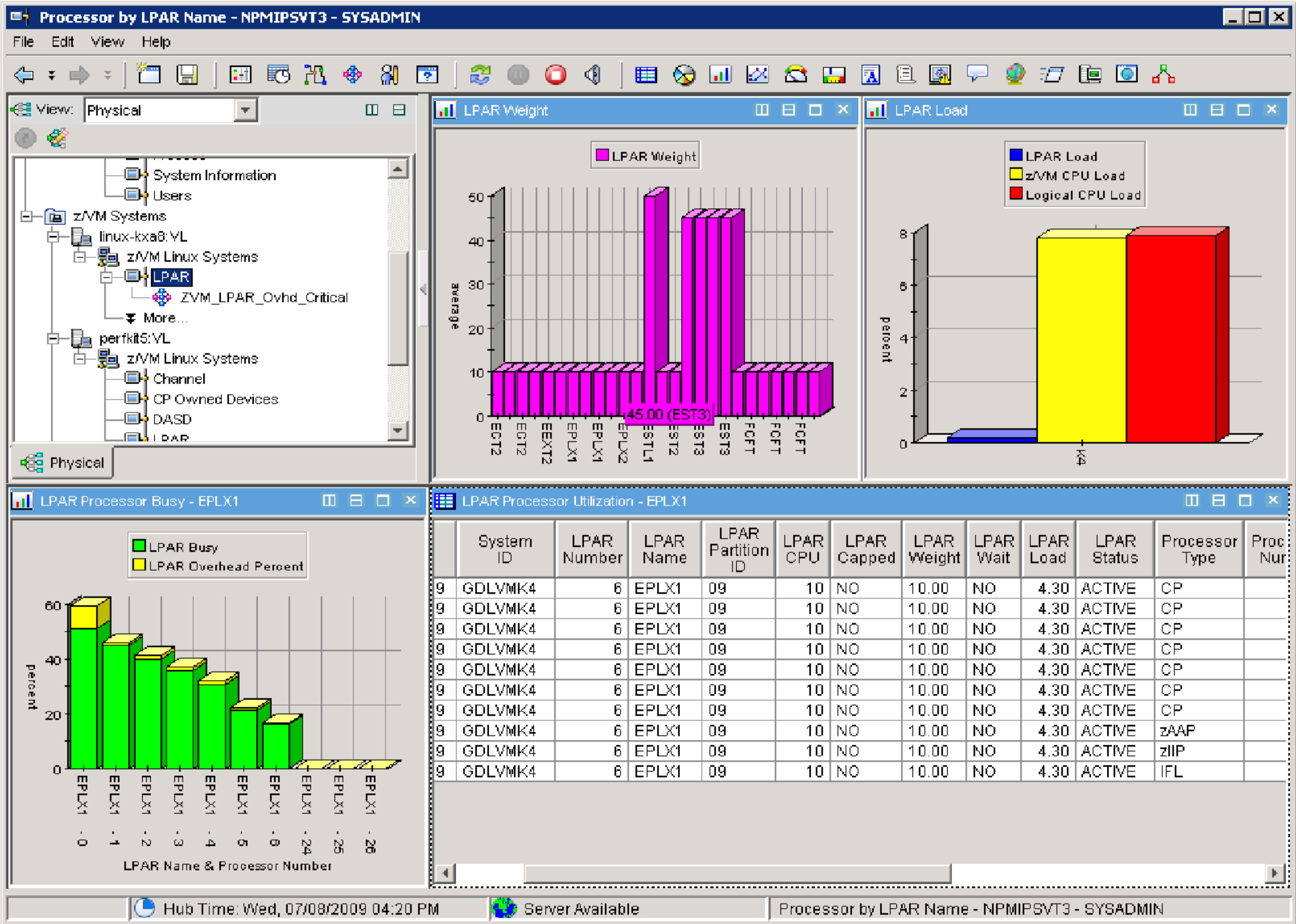


Scenario—CPU Usage

- **Considerations**
  - Are the LPAR weights balanced as you expected
  - Do you have an dedicated processors, and should you?
  - Do you have the right number of processors per LPAR?
    - Too few means tasks must wait for a processor
      - Multi-threaded applications can use multiple processors concurrently
    - Too many means extra overhead
  - Look at the %Busy
    - Are there some processors being underutilized?
      - Are they needed?
      - Can you shift work (virtual machines) to them
  - Look at multiple intervals, or use historical data before jumping to conclusions
  - May want to drill down to individual processors for more details

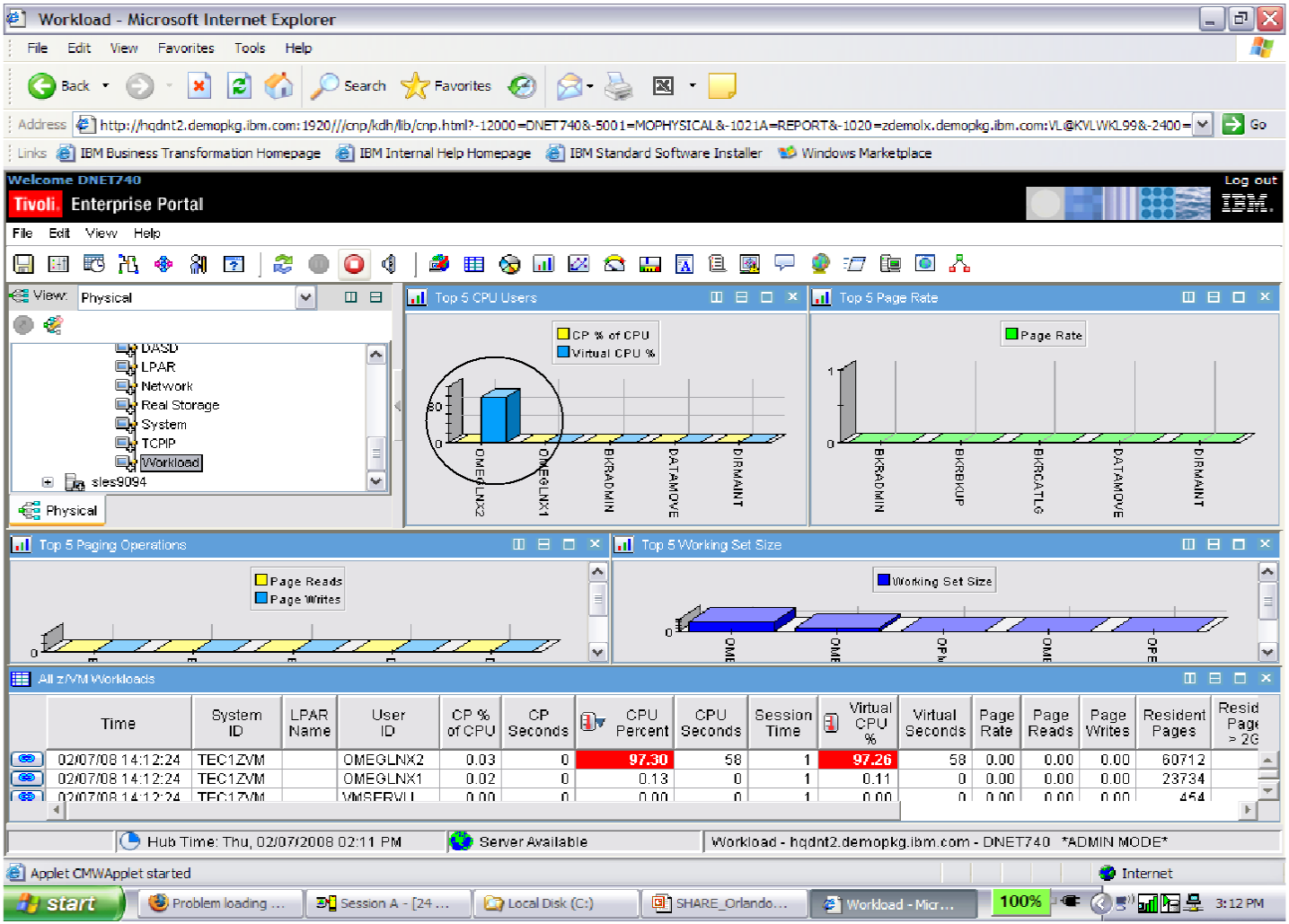
Scenario—CPU Usage—Processor View

Processor workspace allows you to look at all the processors defined for a particular LPAR



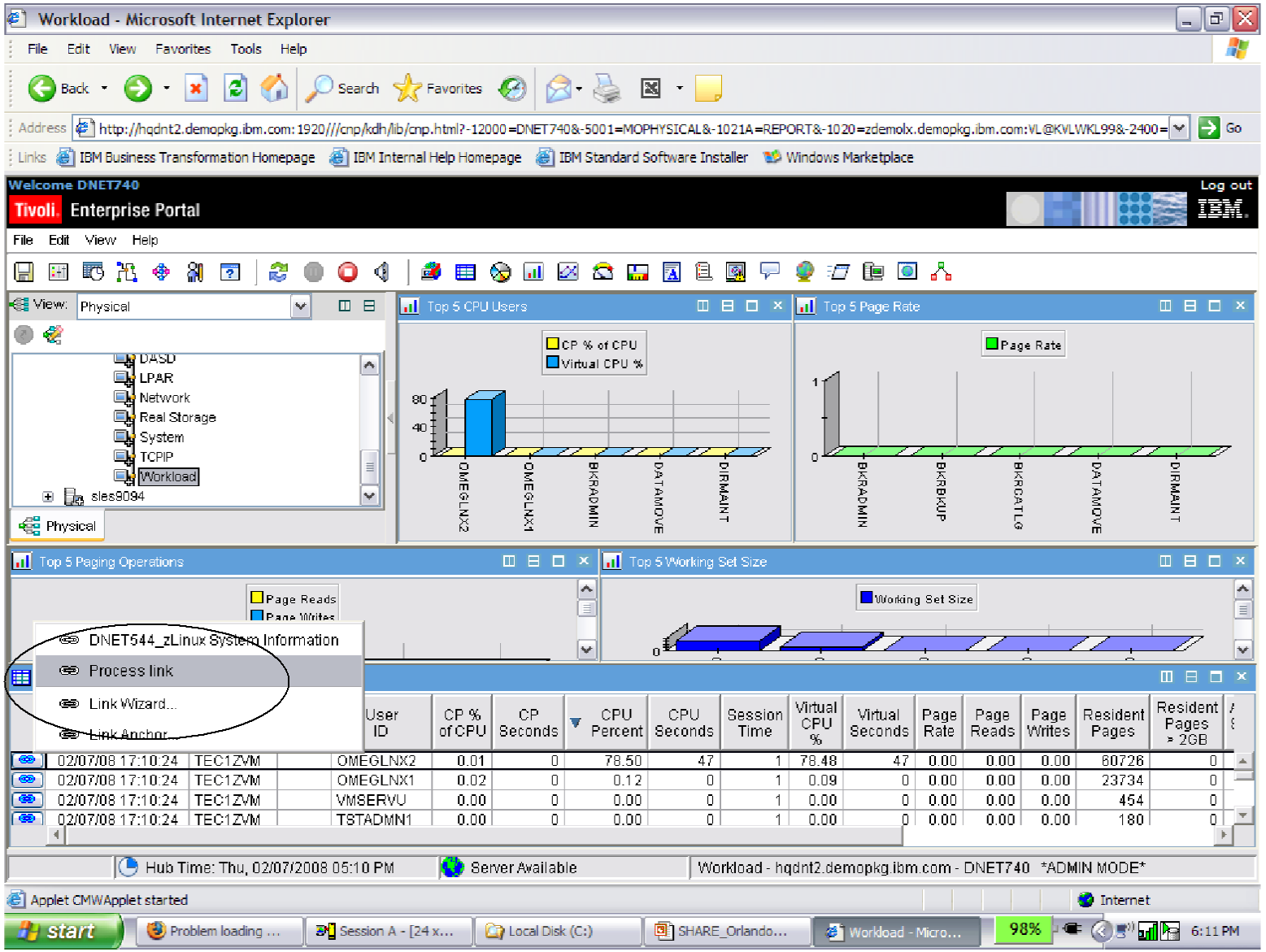
Scenario—System Running Slowly

System is running slowly. Check Workload workspace to see if any particular user is hogging the CPU.



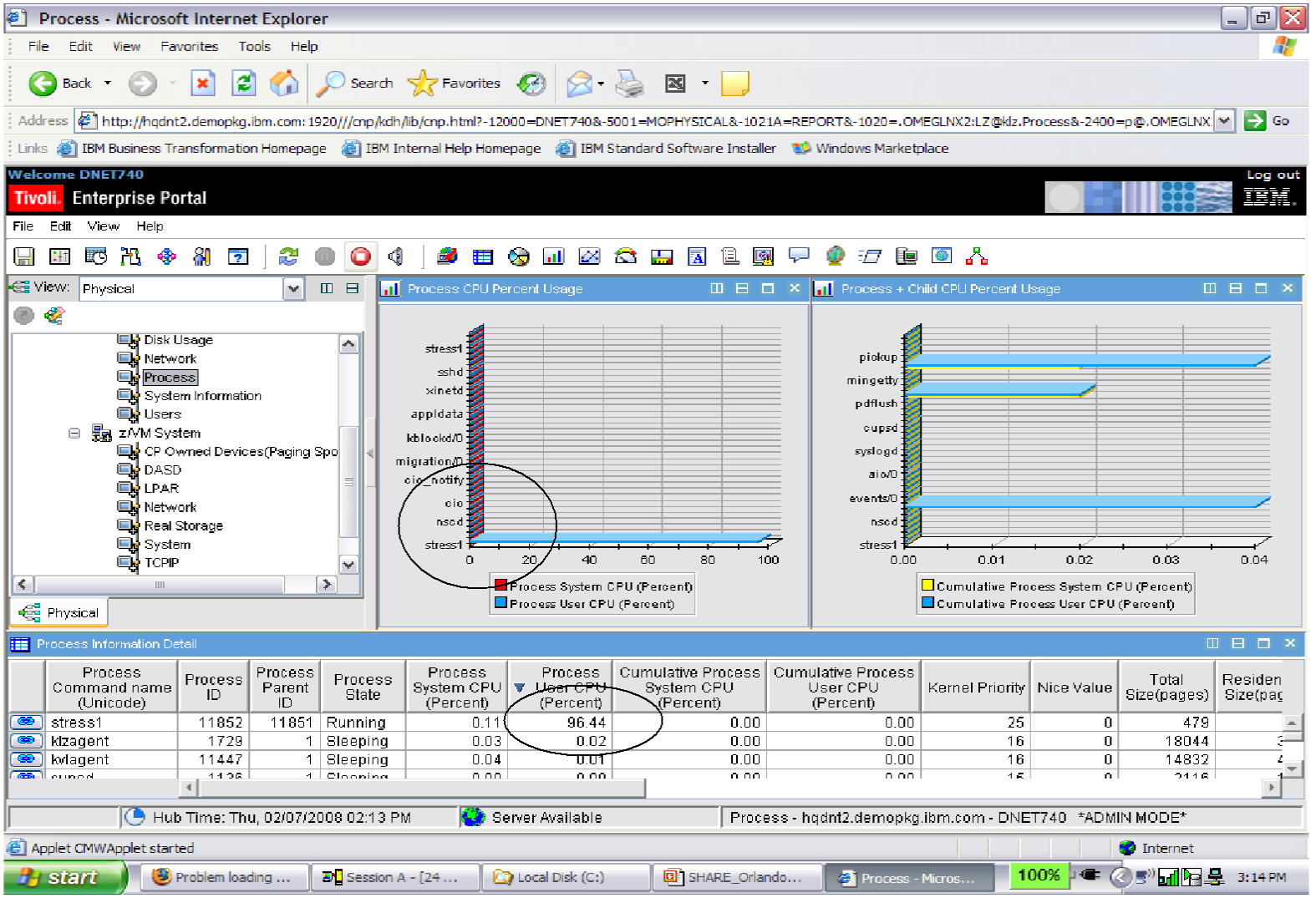
Scenario—System Running Slowly (cont)

Predefined Link to take You directly To the Process workspace



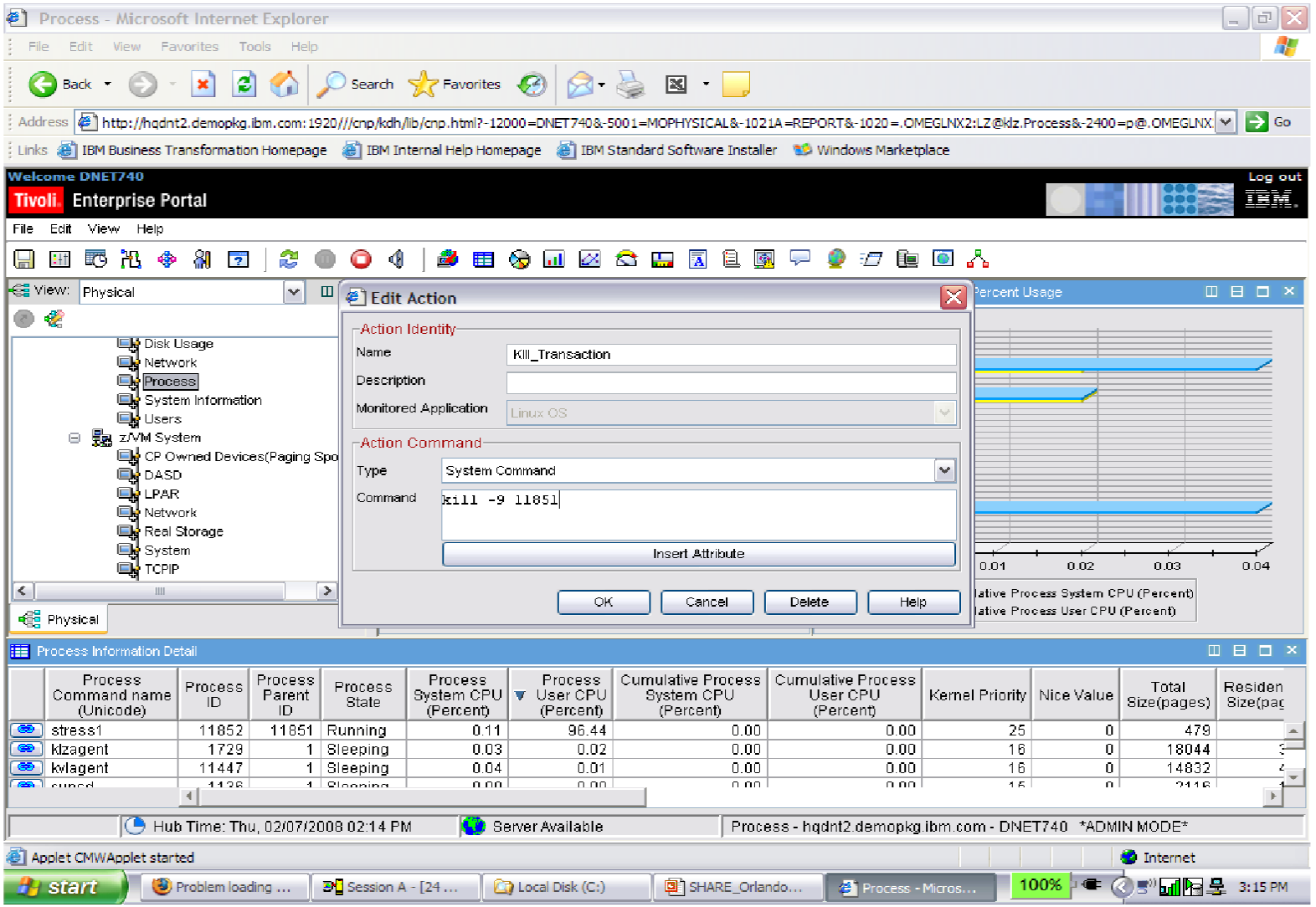
Scenario—System Running Slowly (cont)

See if there is a process which is using too much CPU



Scenario—System Running Slowly (cont)

You can issue a Take Action command to stop the offending process



Why choose IBM for your z/VM and Linux on z management needs ?

- **End to End Management and Seamless integration with other Tivoli Monitoring products through the Tivoli Enterprise Portal**
  - Other solutions have multiple inconsistent user interfaces and cannot provide an end to end view spanning cross platform applications.
  - Some solutions are silo oriented and do not have the breadth and depth to manage the applications that are running on z/VM and Linux
  - If you are considering WebSphere, SAP, Oracle Financials, UDB,Oracle DB, etc for this platform IBM has a fully integrated suite of monitoring tools across distributed and zSeries environments. Other solutions tend to only have consolidated alert consoles, with minimal launch in context capabilities
- **Full suite of z/VM and Linux performance and management tools**

IBM Management Portfolio for z/VM and Linux on z

IBM System z Virtualization Infrastructure
- IBM System z hardware (including LPAR hypervisor)
- IBM z/VM Version 5

Monitoring for Virtualization Infrastructure
- z/VM Virtual Machine Resource Manager (included with z/VM)
- IBM z/VM Performance Toolkit for VM (z/VM priced feature)
- IBM Director
- IBM Tivoli OMEGAMON XE on z/VM and Linux
- IBM Tivoli Monitoring
- IBM Tivoli Composite Application Manager for SOA
- IBM Tivoli Usage and Accounting Manager

Automation for Virtualization Infrastructure
- IBM Operations Manager for z/VM
- IBM Tivoli Enterprise Console
- IBM Tivoli Workload Scheduler

Provisioning Management
- IBM z/VM DirMaint (z/VM priced feature)
- z/VM Center task of IBM Director
- IBM Tivoli Provisioning Manager

Resiliency Management
- IBM Tivoli System Automation for Multiplatforms

Application Layer Management
- IBM Tivoli Application Dependency Discovery Manager
- IBM Tivoli OMEGAMON XE for Messaging
- IBM Tivoli Composite Application Manager for Response Time
- IBM Tivoli Composite Application Manager for Web Resources
- IBM Tivoli Composite Application Manager for Transactions
- IBM Tivoli License Compliance Manager

Extended Infrastructure Management *(Security)*
- IBM z/VM RACF Security Server (z/VM priced feature)
- IBM Tivoli zSecure
- IBM Tivoli Access Manager for e-business
- IBM Tivoli Access Manager for OS
- IBM Tivoli Federated Identity Manager
- IBM Tivoli Identity Manager
- IBM Directory Server
- IBM Directory Integrator
- IBM Tivoli Risk Manager

Extended Infrastructure Management *(Storage)*
- IBM SAN Volume Controller (SVC)
- IBM Tivoli Storage Manager
- IBM TotalStorage Productivity Center
- IBM Backup and Restore Manager for z/VM
- IBM Tape Manager for z/VM
- IBM Archive Manager for z/VM

Extended Infrastructure Management *(Network)*
- IBM z/VM RSCS (z/VM priced feature)

Business Services Management
- IBM Tivoli Business Service Manager
- IBM Tivoli Service Request Manager
- IBM Change and Configuration Management Database (CCMDB)

For specific releases, refer to Tivoli Platform Support Matrix at: **ibm.com**/software/sysmgmt/products/support/Tivoli_Supported_Platforms.html

Demo

Bring it all together

It is often that a unit of work is serviced by multiple applications and databases across multiple operating systems, including z/VM and Linux. Integrated views allow:

- Unit of work, or application tracking
- Business views
- Single skill sets to monitor dissimilar hardware, operating system, and application environments.

Application View: Scaling Scenario

- **WebSphere MQ on Linux for System z receives "order requests" in the form of Queue messages, and places them on a queue.**
- **A WebSphere Application Server is invoked to periodically check the queue for messages and process them to a DB2 on z/OS database.**
- **The orders are coming too fast for the Websphere application to process.**
- **A second Linux server is started with another copy of Websphere application server to aid in the processing of requests.**

Application View: Scaling Scenario

- **Trigger: Queue Depth**
- **Options for triggering actions can be based on things such as:**
  - The number of orders received but not yet processed (the number of messages on the queue)
  - The amount of time it is taking to process the orders
  - The response time of the web application
  - The CPU usage of the z/VM Guest
  - Other things I haven't given much thought to yet.

धन्यवाद
Hindi

多謝
Traditional Chinese

감사합니다
Korean

Спасибо
Russian

Gracias
Spanish

Thank
English

You

شكرا
Arabic

Obrigado
Brazilian Portuguese

Danke
German

多谢
Simplified Chinese

Grazie
Italian

Merci
French

நன்றி
Tamil

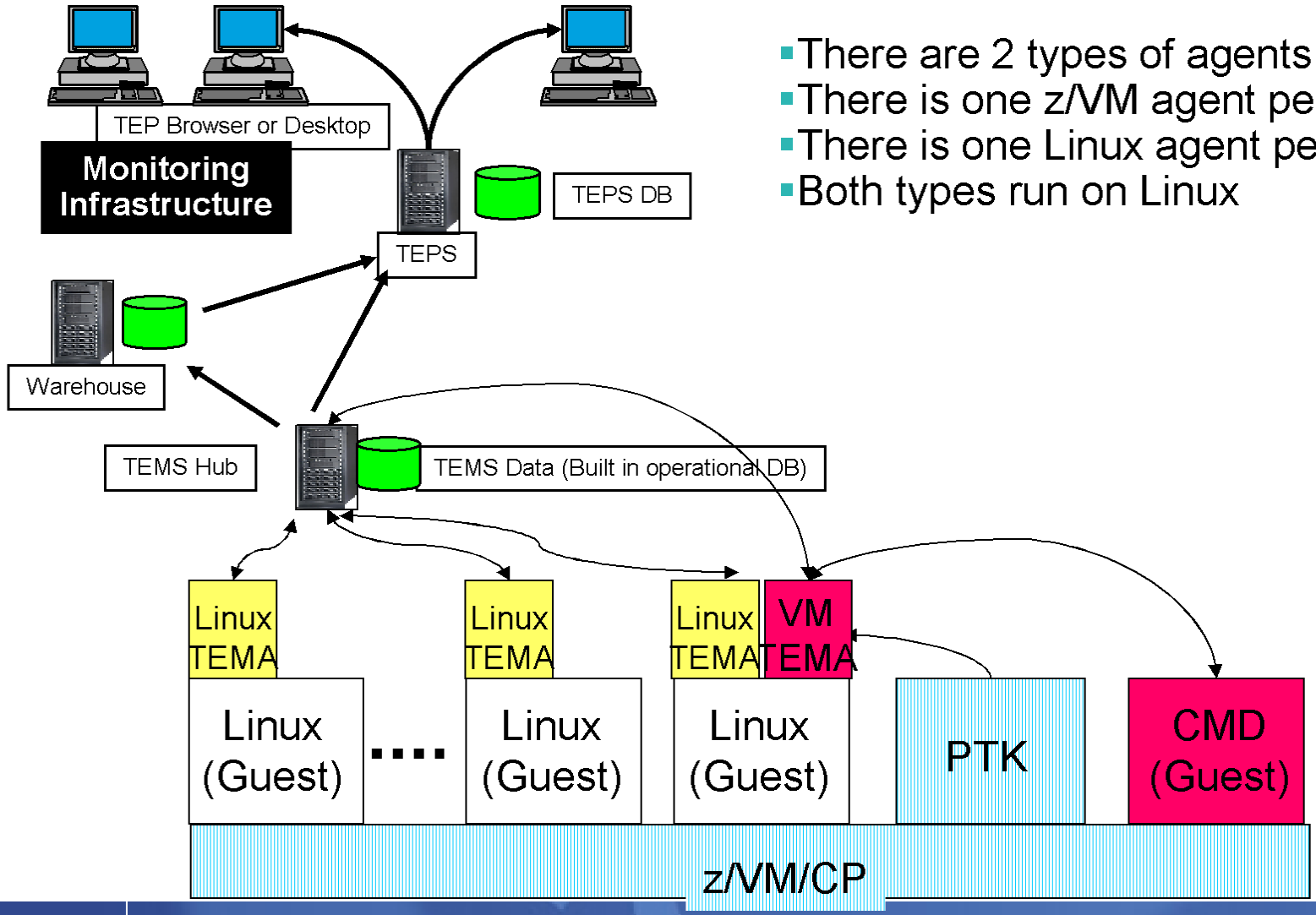
ありがとうございました
Japanese

ขอบคุณ
Thai

IBM Software

Backup Charts

OMEGAMON XE on z/VM and Linux agents



- There are 2 types of agents
- There is one z/VM agent per z/VM LPAR
- There is one Linux agent per Linux Guest
- Both types run on Linux

Architecture

- **OMEGAMON XE on z/VM and Linux:**
  - Provides monitoring across two Operating systems
    - Linux on z, and z/VM
  - Smartly incorporates Performance Toolkit (PTK) to
    - Reduce duplicate monitoring
    - Provide an OMEGAMON II and XE like solution.
  - Requires components on z/VM, Linux, ITM, and the Performance Toolkit program product.
  - z/VM and Linux agents on CD, shipped tape contains Command Processor.

Performance

- **Linux OS agent is very low overhead**
- **Typically .01-.015% of one CPU for a Linux agent running alone on a z/VM guest.**
- **Running situations, and/or collecting history data in addition to regular TEP displays, increases the overhead.**

Key Features in Most Recent Release (V4.1.2)

- **Additional data for Control Unit and DASD Cache, VDISK, Spin Locks.**
- **Additional VSWITCH, LPAR, and Processor data.**
- **V4.1.2 went GA in December 2008**

Prerequisites

- z/VM 5.3 or z/VM 5.4
  - Performance Toolkit at latest service level
- SLES 9 with Service Pack 3, or SLES 10
- RHEL 5 or later
- Need at least one Linux on z image to support the agent.
- Does not have to be a Linux dedicated to the agent.

OMEGAMON XE 4.1.0.3

- **VM64299 – FL520 & FL530**
  - Provides base function in FL520
  - Linux normalization updates in both FL520 and FL530
  - Mixed engines support in FL530

OMEGAMON XE 4.1.1

- **VM64337 - FL520 & FL530**
  - Support minidisk cache
  - Control unit data
  - Channel data

OMEGAMON XE 4.1.2

- **Included in FL540 base**
- **Available in FL530 with APAR VM64500**
- **Provides support for**
  - Channel Cache
  - DASD Cache
  - Spin Locks
  - Virtual disks in storage

Installing OMEGAMON XE for z/VM and Linux

- **Install ITM infrastructure first**
  - TEMS – z/OS, Linux/Unix or Windows
  - TEPS – Linux/Unix or Windows
  - TEPD or browser or Java Web Start – Linux on Intel or Windows
    - If using TEPD do not forget to install application support for ITM and OMEGAMON agents

Checklist

- **On z/VM**
  - Install the correct level of Performance Toolkit
  - Build and save the DCSS
  - Add OPTION APPLMON to all Linux guests and TCPIP servers
  - Update the PROFILE TCPIP
  - Update the FCONX $PROFILE to collect OMEGAMON data
  - Enable the Monitor Domains for the data you wish to collect
  - Configure the command processor for the Take Action command (optional)

Checklist (cont)

- **On Linux guest**
  - Ensure you have installed the required packages
  - Configure Linux storage to access the DCSS
    - Either with a mem= parameter in zipl.conf
    - Or DEF STOR CONFIG and leaving a memory hole for the DCSS
  - Do a “modprobe dcssblk” to load DCSS support
  - Do an “echo perfout > /sys/devices/dcssbld/add to link to the DCSS
  - Do modprobe’s for the 3 Linux appldata modules and then enable them
  - Set the timer interval and enable it by updating the /sys/proc/appldata/interval and /sys/proc/appldata/timer files
  - Enable Dynamic Workspace Linking (DWL)
  - Enable sudo for the Command Processor and Take Action (optional)

Checklist (cont)

- **Installing support files (**this is the most commonly overlooked step**)**

  - Be sure to install the application support files at:
    - TEMS
    - TEPS
    - Each instance of the Desktop Client if you are using it

Thoughtful Data Collection with OMEGAMON

- **For ITM 6.2.1 (and later) Linux agent, disable Watchdog.**
- **Agentless data collection available for Linux.**
- **It is easy to enable history collection for every attribute group, but don't do it unless you have the resources.**
  - History data collection uses processor resources and **disk space**.
  - Enabling Tivoli Data Warehouse (TDW) for a group provides data roll-off (UNIX: ITMCMD HISTORY). This reduces the size of short-term history.
  - Not enabling TDW means you must enable manual roll-off of short-term history data.
- **Consider the collection and warehouse intervals.**
  - Reduce short-term history disk space by warehousing more often.
  - Use capping: KHD_TOTAL_HIST_MAXSIZE.
  - Use the Summarization and Pruning agent to control disk space consumed by TDW.

Demo

Bring it all together

It is often that a unit of work is serviced by multiple applications and databases across multiple operating systems, including z/VM and Linux. Integrated views allow:

- Unit of work, or application tracking
- Business views
- Single skill sets to monitor dissimilar hardware, operating system, and application environments.

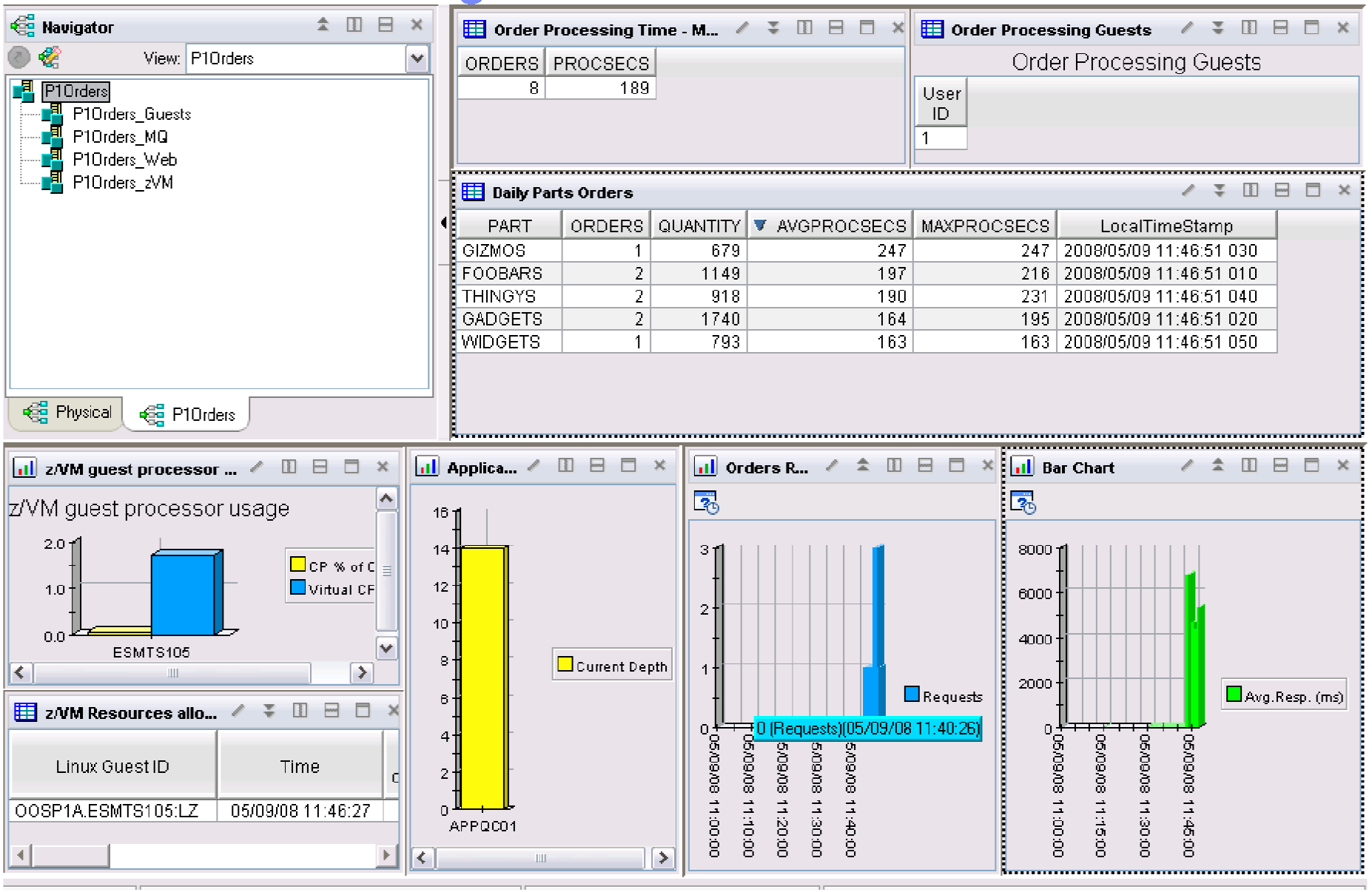
Application View: Scaling Scenario

- **WebSphere MQ on Linux for System z receives "order requests" in the form of Queue messages, and places them on a queue.**
- **A WebSphere Application Server is invoked to periodically check the queue for messages and process them to a DB2 on z/OS database.**
- **The orders are coming too fast for the Websphere application to process.**
- **A second Linux server is started with another copy of Websphere application server to aid in the processing of requests.**

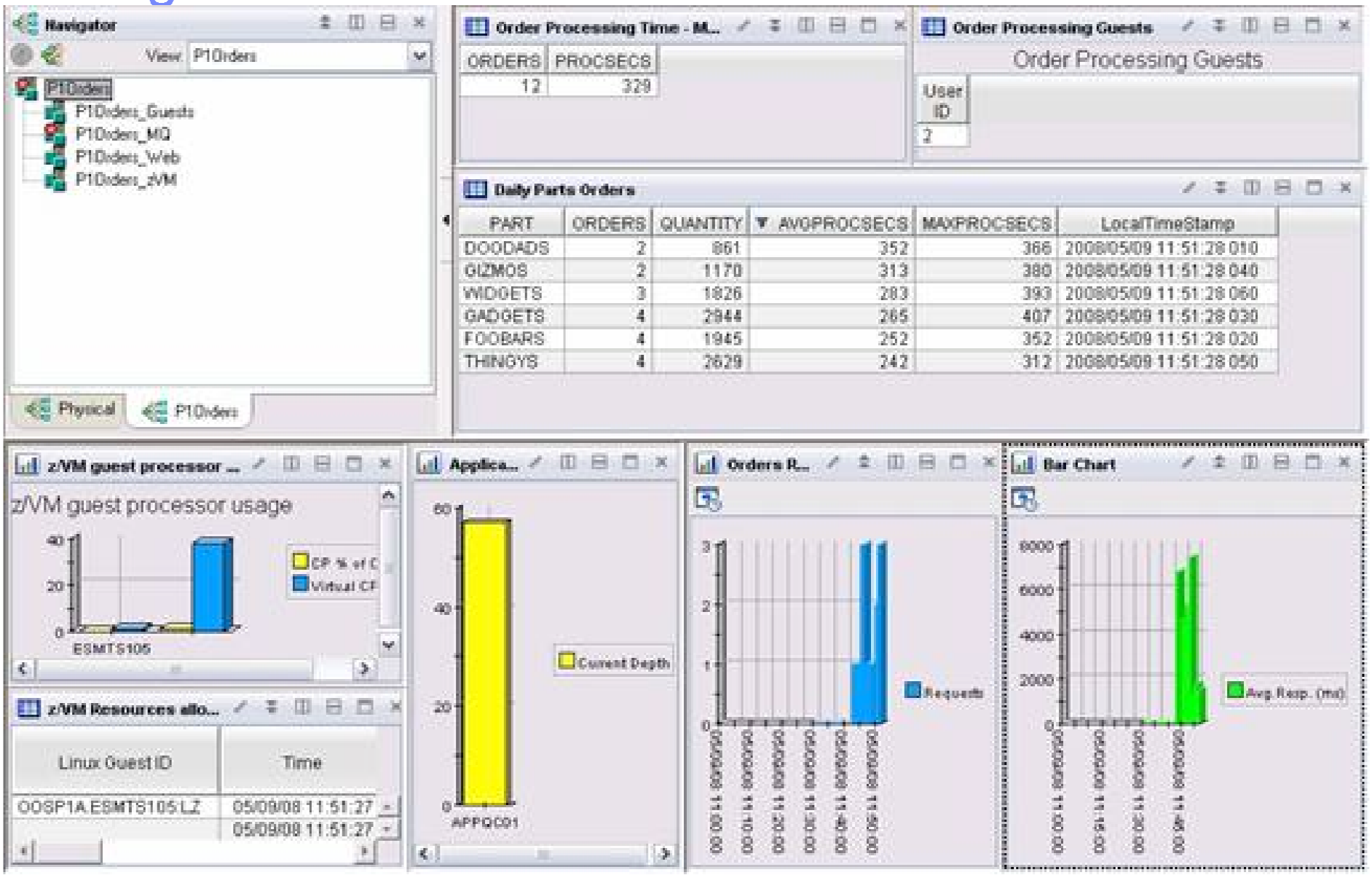
Application View: Scaling Scenario

- **Trigger: Queue Depth**
- **Options for triggering actions can be based on things such as:**
  - The number of orders received but not yet processed (the number of messages on the queue)
  - The amount of time it is taking to process the orders
  - The response time of the web application
  - The CPU usage of the z/VM Guest
  - Other things I haven't given much thought to yet.

MQ Series Queue growth started



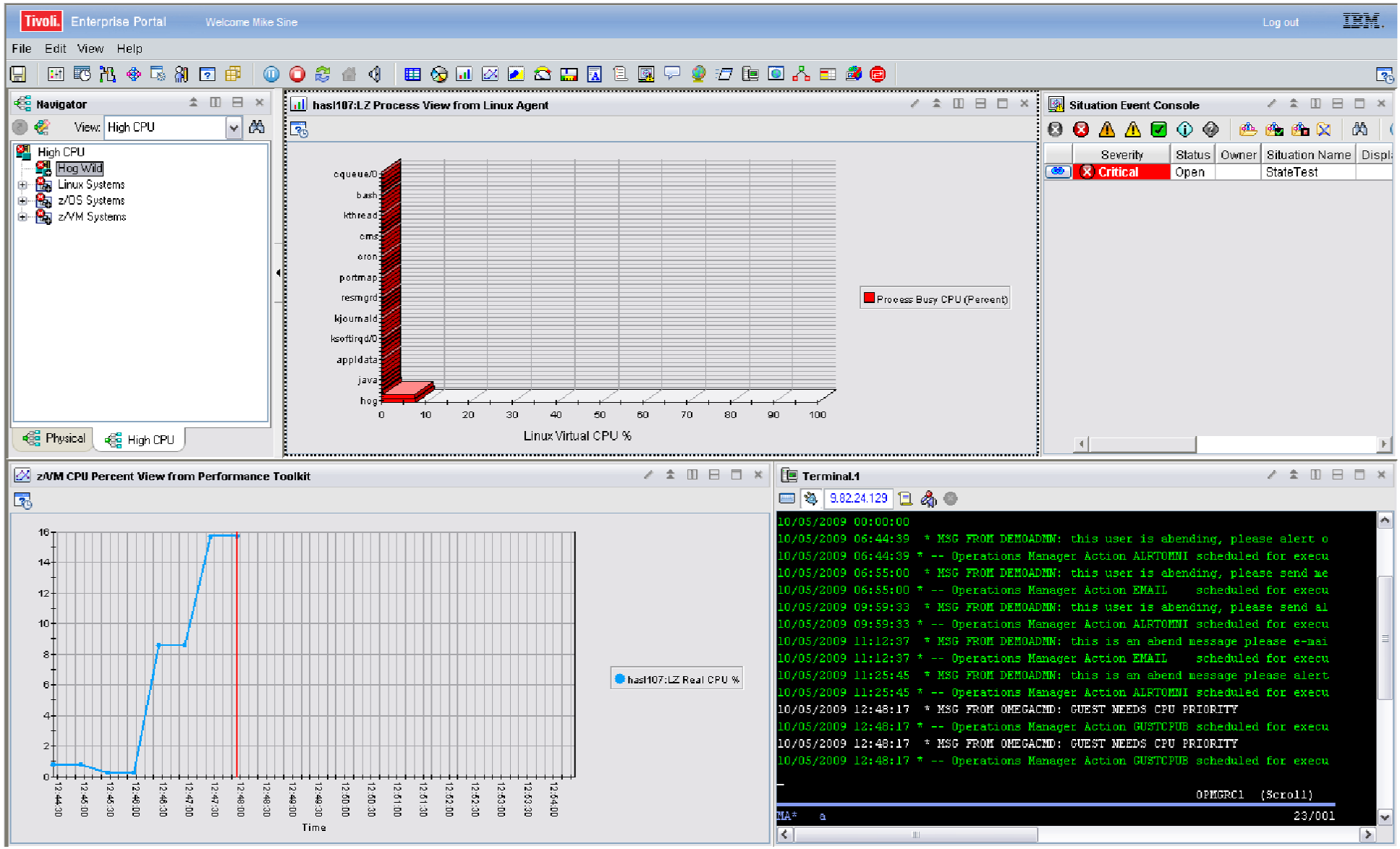
Scaling Scenario



Adjusting Resources for a Linux Guest

- **Virtual CPU consumption is high for a Linux guest**
- **Detect the alert**
  - Automation receives the message
- **Action is triggered by a rule in Operations Manager**
- **Operations Manager issues CP commands to tune the guest**
  - SET QUICKDSP
  - SET SHARE
- **Ability to monitor the output is key**

Adjusting resources for a Linux guest



OMEGAMON Configuration

- **Define a situation (alert) to detect high CPU consumption for Linux virtual machines.**
- **Define the automated "Take Action" to:**
  - Direct a message to console monitored by Operations Manager.
  - Include in the message keywords to trigger Operations Manager rule.
    - Guest Name
    - Guest need CPU priority text
    - Any unique data desired for specific customer environment.