Live Partition Mobility @ Australia Post

2009 IBM Power Systems Technical Symposium. 10-13th August, Sydney, Australia

AUSTRALIA POST

Who is this bloke!?

- Chris Gibson - chris.gibson@auspost.com.au
- AIX Specialist @ Australia Post.
- IBM CATE, System p platform and AIX 5L, technical writer for IBM Systems Magazine, IBM developerWorks and a co-author of the IBM Redbooks publication, "NIM from A to Z in AIX 5L."

Purpose

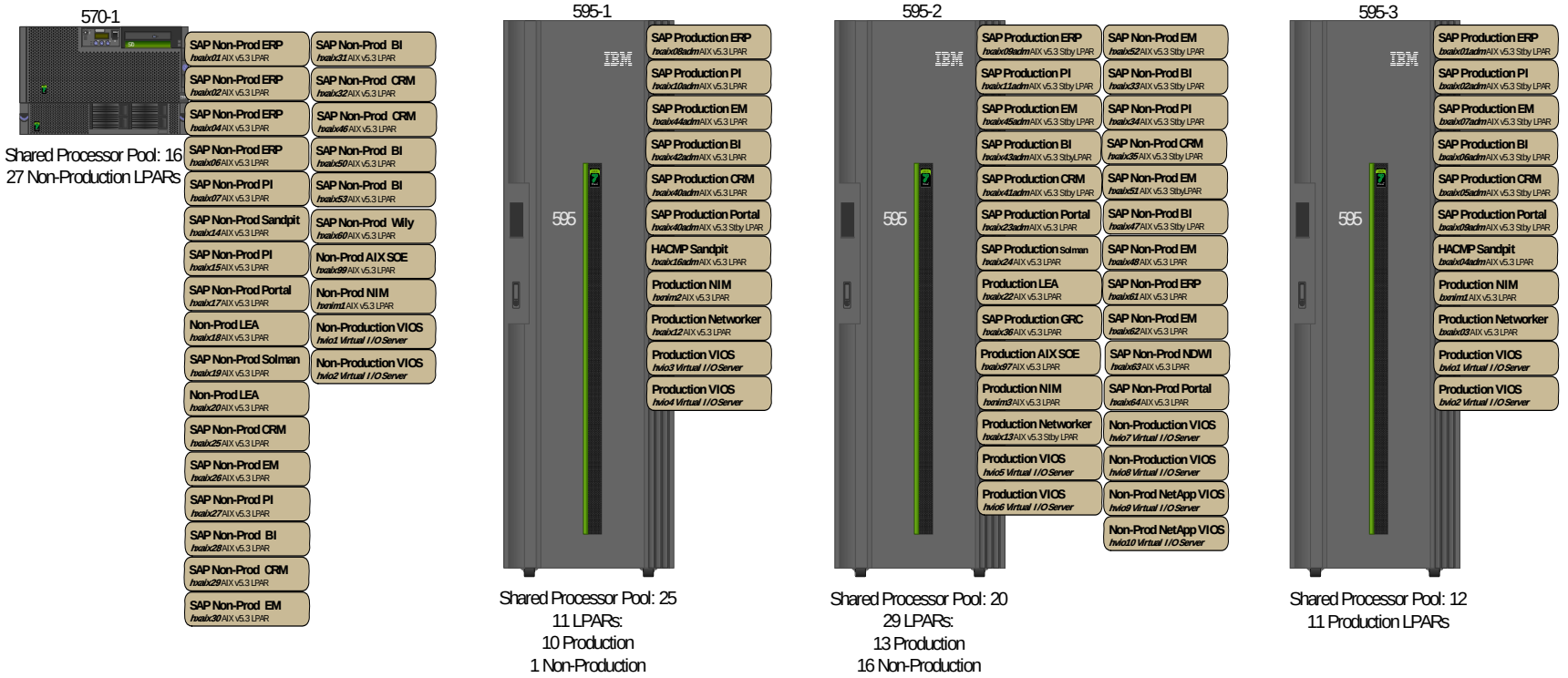
- Share our LPM experience with AIX community.
- Discuss configuring LPM on JS22 and 570/595 systems.
- Exchange ideas with other AIX customers.
- Demonstrate use of latest technology, outside of IBM.
- Provide feedback to IBM development.

Audience

- Technical
- AIX
- SAP

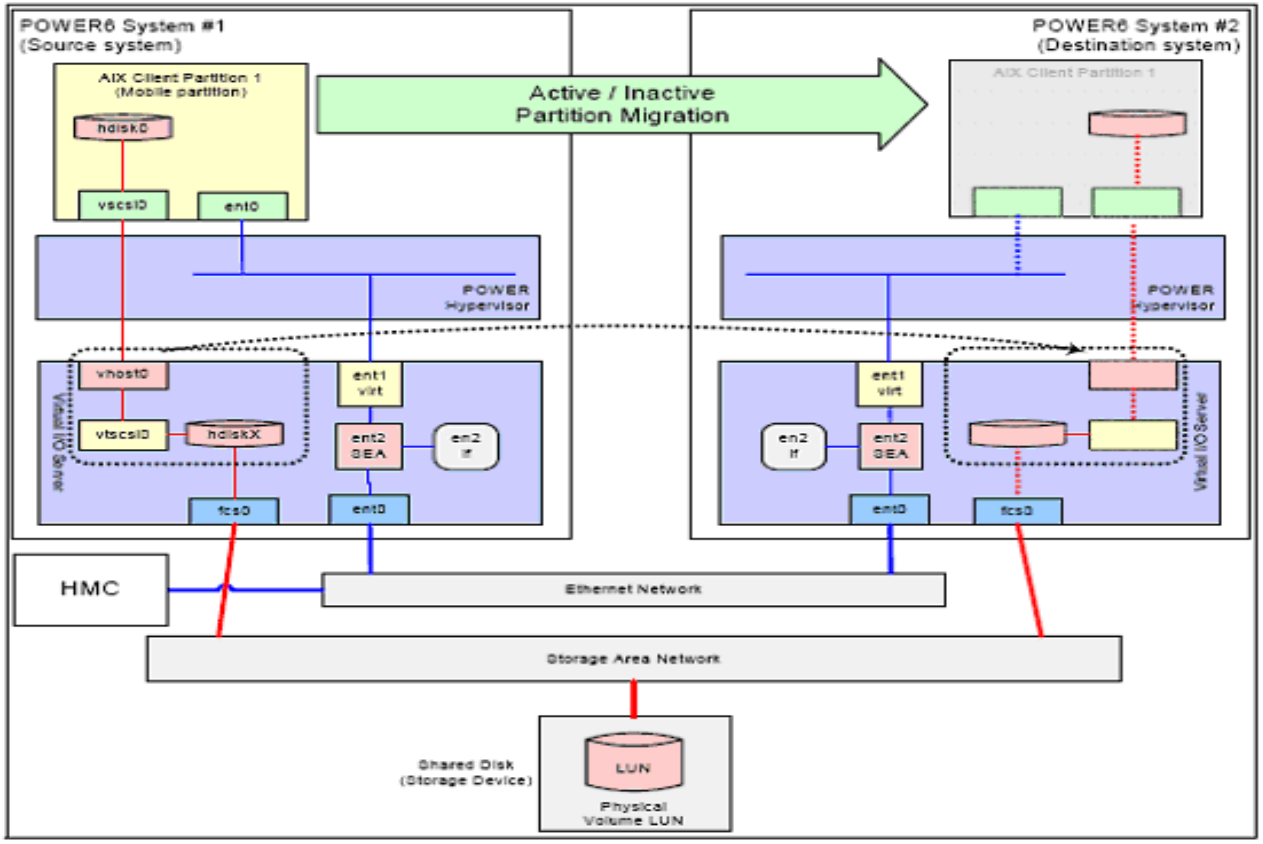
AIX-POWER6 Landscape @ POST

- ~80 LPARs



Partition Mobility in brief...

- Live Partition Mobility (LPM) allows you to move an active LPAR from one physical server to another without disrupting users or running processes. The migration transfers the entire LPAR (Logical Partition) state, including processor context, memory, connected users, running batch jobs etc. Attend LPM sessions at this event to learn more!



Which systems? Benefits of LPM?

- Live Partition Mobility is available on POWER6 based systems.
- It enables the migration of an active LPAR from one physical system to another. LPARs must be virtualised i.e. shared processor and VIO.
- Mobility uses a simple procedure that transfers the LPAR from the source to the destination system without disrupting the hosted application or the operating system.
- It allows an administrator to perform hardware maintenance, such as disruptive firmware updates, without requiring system downtime. LPARs can be temporarily moved to different physical servers during the maintenance window. They can be easily moved back once the outage is complete.
- It provides an administrator greater control over the usage of System p resources as workload can be moved dynamically between systems.
- Live Partition Mobility is targeted for planned activity. It does not protect you from system failures, so it does not replace high-availability software like the IBM HACMP (PowerHA) high-availability cluster technology.

Australia Posts direction for LPM.

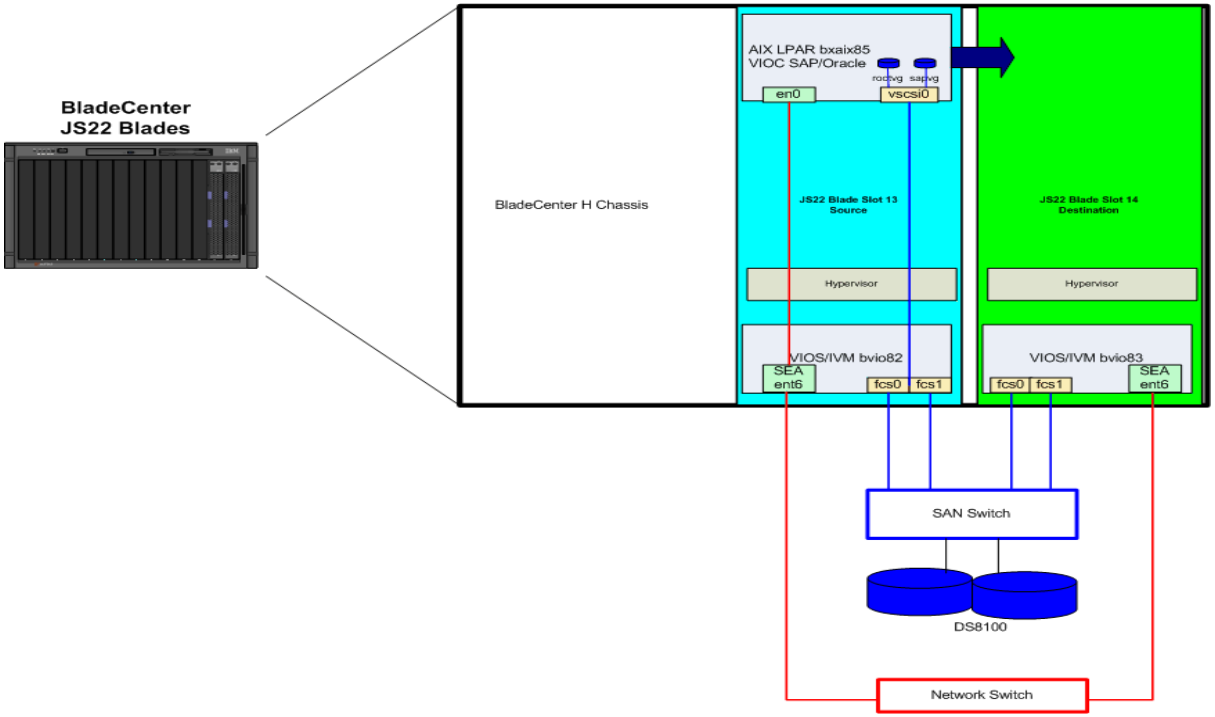
❑ LPM is the latest weapon that Australia Post will use on the road to continuous availability. It will help to:

- Reduce planned down time by dynamically moving running LPAR from one server to another.
- React to increased workloads over month end by moving non essential LPARs from heavily loaded servers to less used servers – leaving spare capacity for critical month end workload.
- Develop an energy reduction program which allows easy consolidation of LPARs.

❑ *But* how do we now it works in our environment? i.e. with our SAP applications? Need a POC project to develop and test.

Proof of Concept Environment

- Prove concept in non-production lab. Prior to *real* non-prod and Production.
- Two JS22 Blades. Each running VIOS 2.1 and IVM. 16GB Memory and 4 x POWER6 processors.
- One LPAR running AP AIX SOE v6.1 TL2 SP2*.
- One SAP R3 4.7 instance with Oracle 10G.



*When first tested, our levels were VIOS 1.5.1.1-FP-10.1 and AIX 5.3 TL7 SP3!

Non-Prod Live Partition Mobility@POST

- POC would assist us defining the requirements, process and execution of Live Partition Mobility across the Australia Post non-production IBM AIX POWER6 systems e.g.
  - Gathering requirements for LPM of a non-production LPAR.
  - Raising Remedy and Change Requests for LPM.
  - Executing an LPM operation in the non-production AIX environment at Australia Post.

- Out of scope:
  - LPM operations to and from our DR site. All managed systems must be connected to the same HMC.
  - LPM for production SAP/Oracle systems. Formal SAP testing with Oracle is on-going. Support announced at a later date (please do not attempt to perform a live partition migration of a production Oracle database partition before the formal support is announced). Refer to SAP Note 1102760*.
  - LPM for HACMP clustered systems. HACMP is supported with LPM but has not been thoroughly tested within Australia Posts environment. Further validation required for HACMP and LPM.
  - http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/FLASH10640

*Has since been announced!

POC Objective

- I wanted to be able to move an active AIX LPAR, running SAP and Oracle, from one JS22 Blade to another physical Blade.
- This would provide me with the ability to perform disruptive hardware and/or software maintenance on a Blade without the need for an outage to my SAP applications.
- For example, if I had a requirement to upgrade a VIOS on a Blade, I could move the workload to another Blade (without an outage to SAP) , perform the VIOS update, reboot the Blade, and then move the LPAR back once the activity was successfully completed.
- Likewise, I could take the same action if I needed to update the Blade's firmware.
- Blades used for Pilot and POC SAP applications. 1-2 users.
- Once proven we could plan for and verify LPM across the 570 and 595 landscape.

JS22 Blade Environment

- IBM BladeCenter H chassis.
- JS22 Blades in slots 13 and 14, respectively.
- Both Blades have 16GB of memory installed, 4 x 4GHz POWER6 processors
- 'PowerVM Enterprise Edition‘ enabled (required for mobility).
- Each Blade was installed with a Virtual I/O server (VIOS, version 1.5*) and Integrated Virtualization Manager (IVM).
- The SAN disk storage for these systems was an IBM DS8100.
- VIOS hostnames: bvio82 (slot 13) and bvio83 (slot 14).
- The Blade in slot 13 had one AIX LPAR (bxaix85) configured and active.
- The application hosted on this system was a single SAP R3 v4.7 instance with Oracle 10G. SAP was installed and configured by our SAP Basis team.
- It is important to note that they did not have to do anything special with their SAP (or Oracle) installation to support the mobility feature*.

*Note Official Support Statement from Oracle.

JS22 Blade Environment – continued.

- There are several prerequisites for partition mobility.
- All network connectivity from the LPAR must be virtualized, meaning it must communicate using a VIOS.
- This implies that the VIOS must have a Shared Ethernet Adapter (SEA) configured and operational.
- Both of my VIOS were configured with an SEA, on the same physical VLAN. I used one of the Logical Host Ethernet (LHE) ports to configure the SEA.
- All of the SEA configuration was performed using the IVM and was very straightforward. No need for *mkvdev -sea*, was a nice change!
- The Virtual I/O Client (VIOC), bxaix85, was configured with a virtual ethernet interface configured with the appropriate VLAN ID to communicate with the outside world using the SEA in the VIOS.

JS22 Blade Environment – continued.

- Another important prerequisite for partition mobility.
- All storage connected to the mobile LPAR must be on the SAN.
- Even the operating system i.e. rootvg.
- SAN disk must be assigned to both Blades and be detected by both VIOS.
- This is to allow the target VIOS the ability to "take over" the storage during a migration.
- I allocated two SAN (DS8100) disks to both VIOS.
- One disk was for the OS (AIX *rootvg*)
- The other was for the SAP/Oracle software and database (*sapvg*).

Configuring the JS22 environment for partition mobility.

- First step. Install Virtual I/O Server (VIOS) on each JS22.
- Accomplished by installing a VIOS mksysb image using NIM.
- Internal disk within the Blade can be used to house the VIO server.
- May choose to boot the Blade with the SAN, as this is also supported.
- I chose the internal disk for my Blades.
- Connect to the Web-based IVM. "HMC-like" GUI.
- IVM allows administrator to configure LPARs, virtual network, and virtual storage on the Blade and VIOS.
- Web-based tool. Simply point your Web browser at the VIOS hostname.
- http://bvio82. Presented with the IVM login page.
- Use the VIOS *padmin* userid and password.

Configuration - continued.

- Update the firmware levels of the JS22 and associated components such as the Fibre Channel (FC) adapters.
- Download the latest firmware images for the JS22 and the FC adapters from the JS22 support site.
- Install the latest VIOS fix pack.
- During the build of my VIOS, the latest fixpack was 1.5.1.1-FP-10.1.
- Install and update multipath I/O (MPIO) device driver. When connecting to an IBM DS8100 storage device, the supported MPIO software is *SDDPCM* v2.2.0.
- With the correct software and firmware levels installed, prepare the Blade, the VIOS, and the LPAR for partition mobility.
- Brief checklist of the tasks performed with the IVM:

1. Enter the PowerVM Enterprise Edition APV key on both Blades. This key is required to enable the mobility feature on the JS22 Blade.

2. Confirm that the memory region size is the same on both Blades. This information can be found under "View/Modify System Properties," in the "Memory" tab.

Configuration - continued.

3. Configure an SEA on both VIOS. Enable the Host Ethernet Adapter for ethernet "bridging".
4. Required for the virtual ethernet devices to access the physical ethernet adapter and the external network.
5. Performed under the "*View/Modify Host Ethernet Adapter*", "*Properties*" tab.
6. Message appears stating the operation was successful. The SEA is now configured.
7. Create an LPAR (bxaix85) on the source Blade.
8. Ensure that none of the *physical* HEA ports are selected. Under "Virtual Ethernet," select the SEA to use (for instance, ent0).
9. Under "Storage Type", select Assign existing virtual disks and physical volumes. Select the *SAN disk* assigned to the VIOS i.e. the DS8100 disks.
10. Next step is to install AIX. Achieved using a NIM mksysb install of AP SOE image.
11. Configure SAP and Oracle.

Configuration - continued.

- Review and check. Each VIOS has an SEA. ent6: LHE port, ent0.

```
$ lsmap -all -net | grep -p ent6
SEA                     ent6
Backing device          ent0
Status                  Available
Physloc                 U78A5.001.WIH076E-P1-T6

$ entstat -all ent6 | grep PVID
PVID:  1        VIDs:  None
```

- Use *lspv* to check that both VIOS have the same PVID associated with the SAN storage (hdisk1, 2, and 3).

```
gibsonc@hxnim2 /home/gibsonc $ dsh -wbvio82,bvio83 ioscli lspv | dshbak
HOST: bvio82
-------------
NAME            PVID                                VG              STATUS
hdisk0          000071dacdc1fe09                    rootvg          active
hdisk1          000071fabe8efb25                    None
hdisk2          000071fabe8f0837                    None
hdisk3          000071dadd6c7a7d                    None

HOST: bvio83
-------------
NAME            PVID                                VG              STATUS
hdisk0          000071facb8888a2                    rootvg          active
hdisk1          000071fabe8efb25                    None
hdisk2          000071fabe8f0837                    None
hdisk3          000071dadd6c7a7d                    None
```

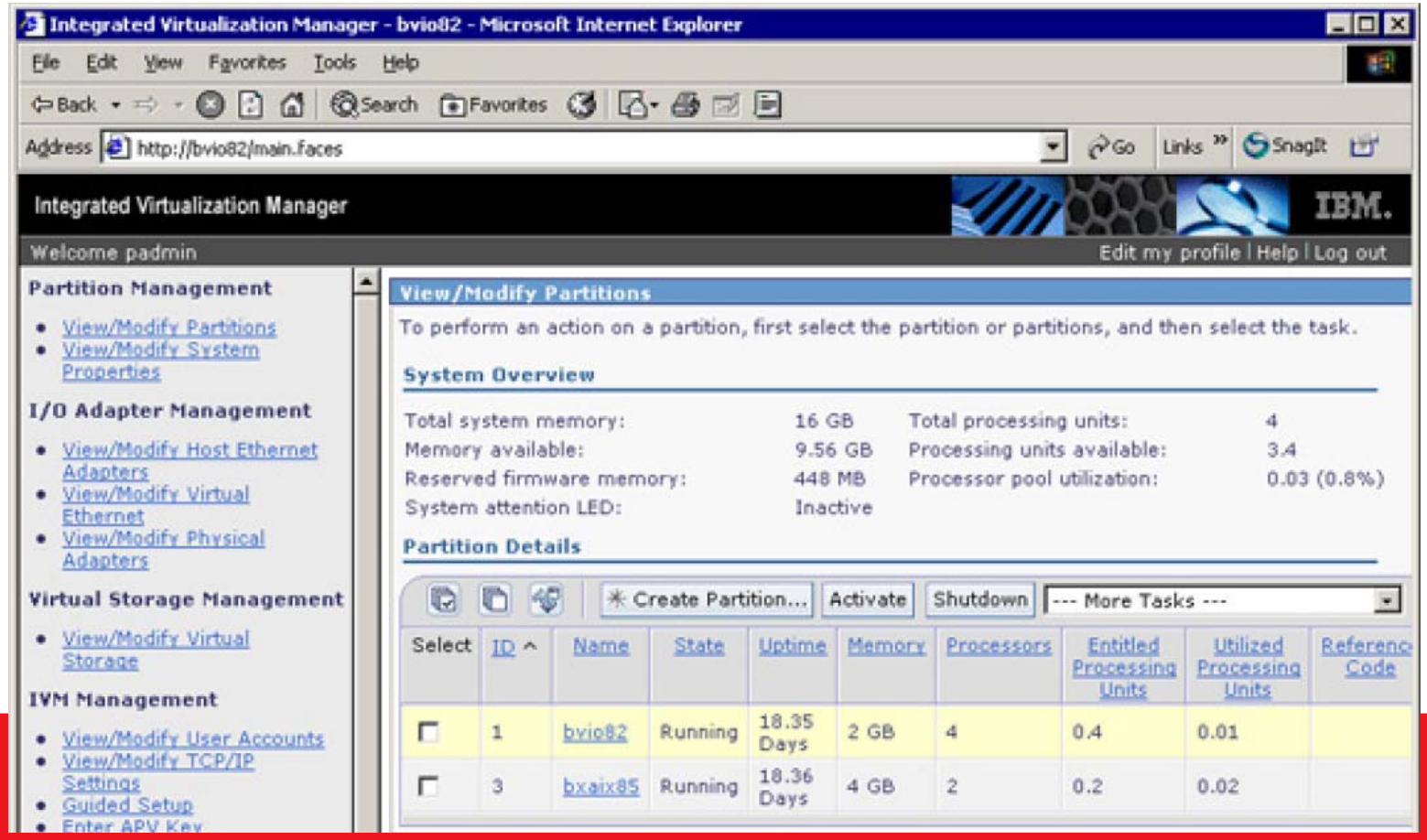
Configuration - continued.

- Ensure that MPIO for the disks is configured and functioning appropriately.
- Run the ***pcmpath*** command (from *oem_setup_env*)
- Verify that all paths are operating normally on both VIOS.
- Confirm that the AIX LPAR, bxaix85, is configured with only virtual devices (meaning no physical adapters, another prerequisite for mobility).
- LPAR is configured with virtual Ethernet and virtual SCSI adapters only.

```
gibsonc@bxaix85 /home/gibsonc $ lsdev -Cc adapter
ent0   Available  Virtual I/O Ethernet Adapter (l-lan)
vsa0   Available  LPAR Virtual Serial Adapter
vscsi0 Available  Virtual SCSI Client Adapter
```

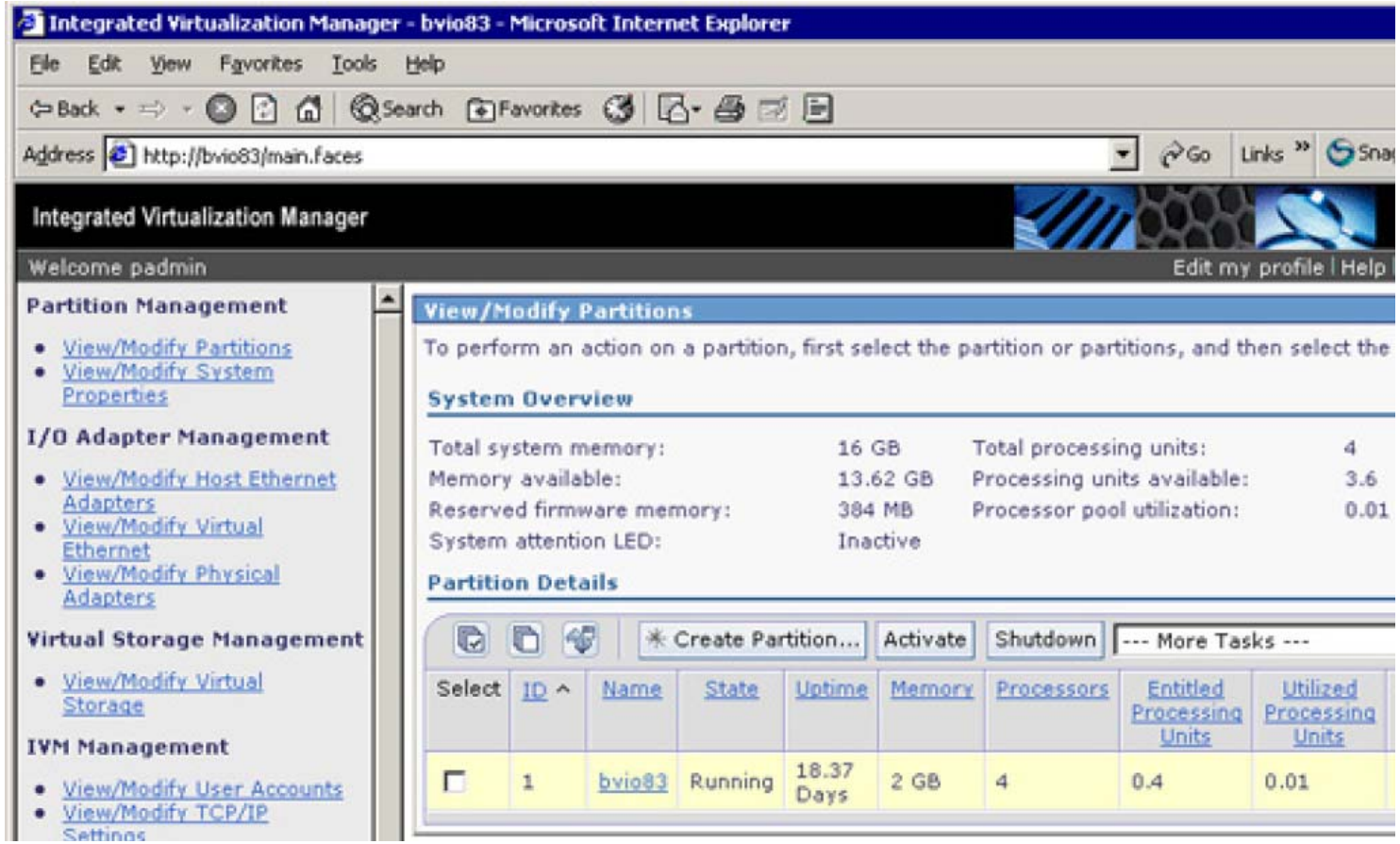
Performing Live Partition Mobility.

- At this point, two VIOS have been configured, bvio82 and bvio83, one per Blade.
- One active AIX LPAR (bxaix85) running on the first Blade as a VIO client (VIOC).
- Ready to perform a live partition migration.
- First Blade (in slot 13, bvio82) known as *source* system.



Performing Live Partition Mobility - cont.

- Second Blade (slot 14, bvio83) the *target* system.



Performing Live Partition Mobility - cont.

- Objective here is to move the LPAR, bxaix85, from the Blade in slot 13 to the Blade in slot 14.
- At the end of the migration, bxaix85 will be running as a VIOC from bvio83 on the other physical Blade.
- AIX, SAP, and Oracle will continue to function throughout the entire migration.
- Prior to the migration, run the *lsconf* command from AIX, and note the system serial number:

```
gibsonc@bxaix85 /home/gibsonc $ lsconf | grep 'Serial Number'
Machine Serial Number: 10071DA
```

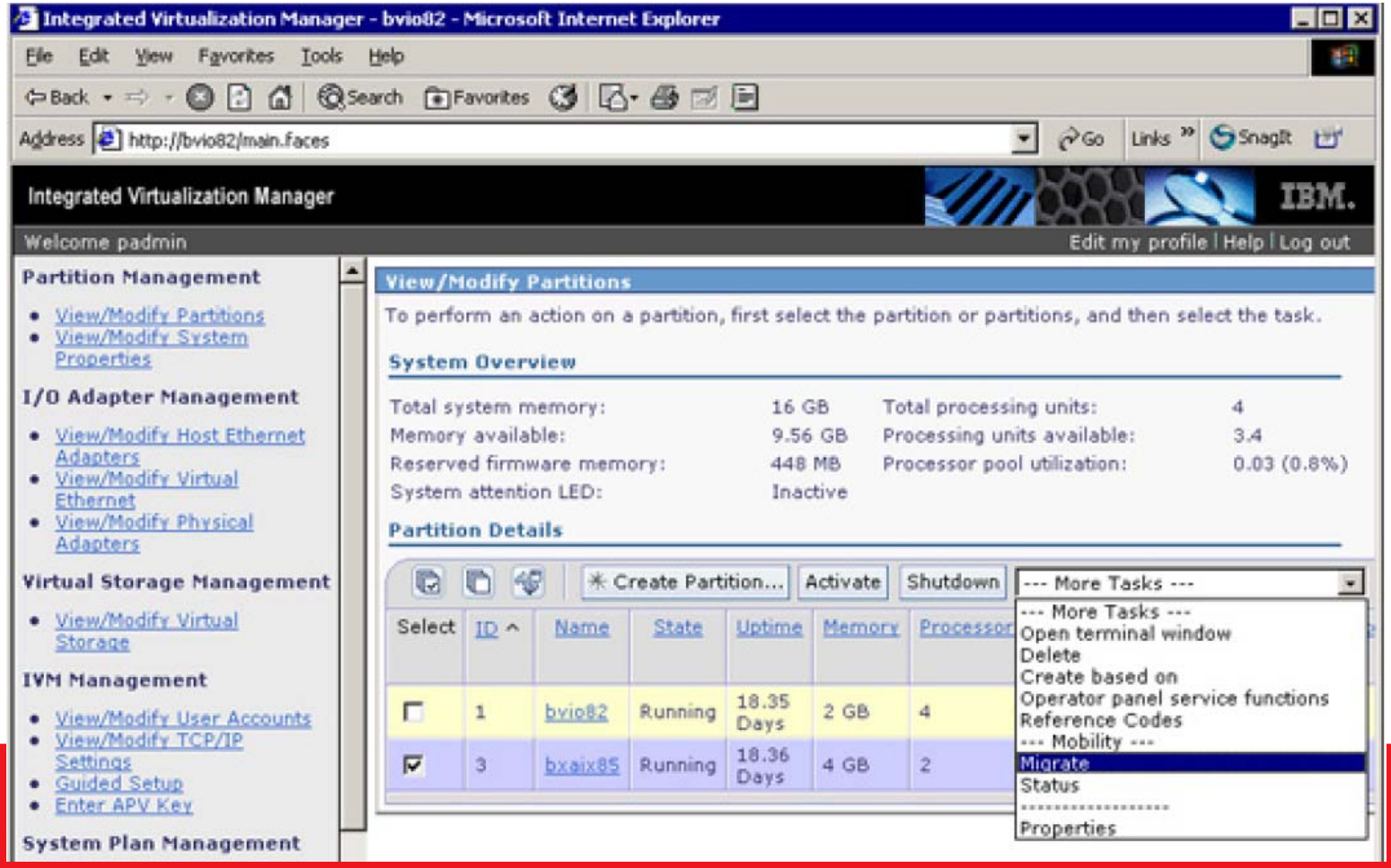
Performing Live Partition Mobility - cont.

- During the migration, there are SAP jobs running on the LPAR.
- Monitor the system using the *topas*
- SAP (disp+work) and Oracle processes are consuming processor during the migration. Running an *SGEN*.

```
Topas Monitor for host:    bxaix85
Thu Sep 18 13:48:06 2008   Interval:  2

Kernel    19.7    |######                              |
User      75.7    |########################            |
Wait       0.1    |#                                   |
Idle       4.5    |##                                  |
Physc =   0.98                        %Entc= 488.8

Network  KBPS    I-Pack  O-Pack   KB-In  KB-Out
lo0    3689.6    6120.5  6120.5  1844.8  1844.8
en0       1.0       2.5     2.0     0.1     0.9

Disk    Busy%      KBPS      TPS KB-Read KB-Writ
hdisk0    2.5     818.0     82.5   444.0   374.0
hdisk2    0.0       0.0      0.0     0.0     0.0
hdisk1    0.0       0.0      0.0     0.0     0.0

Name            PID  CPU%  PgSp Owner
disp+wor     725172  67.3  34.8 cg1adm
oracle       643170  21.6  14.1 oracg1
topas        475228   0.8   2.2 gibsonc
xmwlm        278664   0.2   0.9 root
aioserve     421908   0.2   0.4 root
aioserve     458884   0.2   0.4 root

EVENTS/QUEUES    FILE/TTY
Cswitch    12312  Readch  1945.9K
Syscall    32659  Writech 1880.8K
Reads       6178  Rawin         0
Writes      6134  Ttyout      681
Forks          0  Igets         0
Execs          0  Namei        95
Runqueue     2.0  Dirblk        0
Waitqueue    0.0

PAGING           MEMORY
Faults     3996  Real,MB    4096
Steals        0  % Comp     98.8
PgspIn        0  % Noncomp   1.1
PgspOut       0  % Client    1.1
PageIn       14
PageOut      70  PAGING SPACE
Sios         84  Size,MB   20480
                 % Used      3.3
NFS (calls/sec)  % Free     97.7
ServerV2       0
ClientV2       0   Press:
ServerV3       0   "h" for help
ClientV3       0   "q" to quit
```

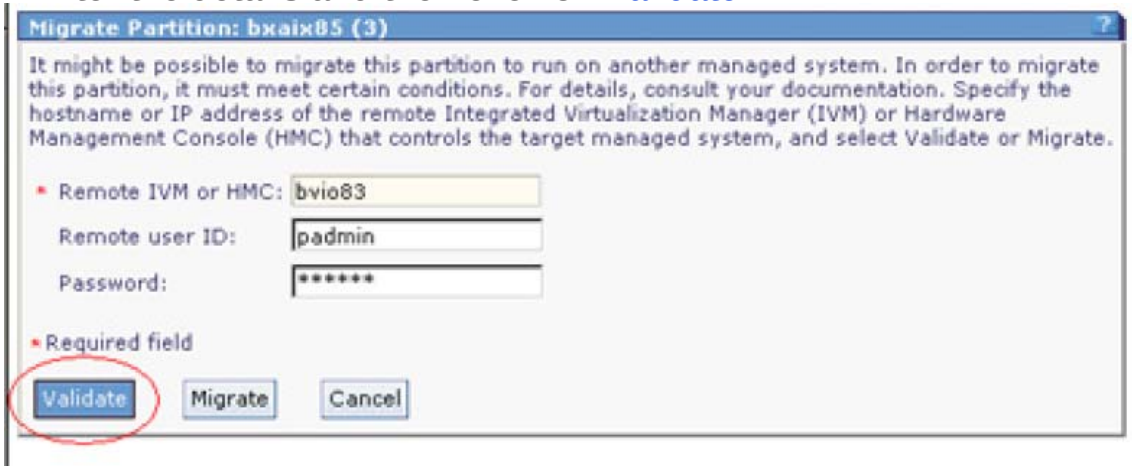
Performing Live Partition Mobility - cont.

- All tasks to perform partition mobility will be executed from the IVM, on the source Blade.
- To start the migration, check the box next to the LPAR (bxaix85).
- Choose Migrate from the "*More Tasks*" drop-down menu.



Performing Live Partition Mobility - cont.

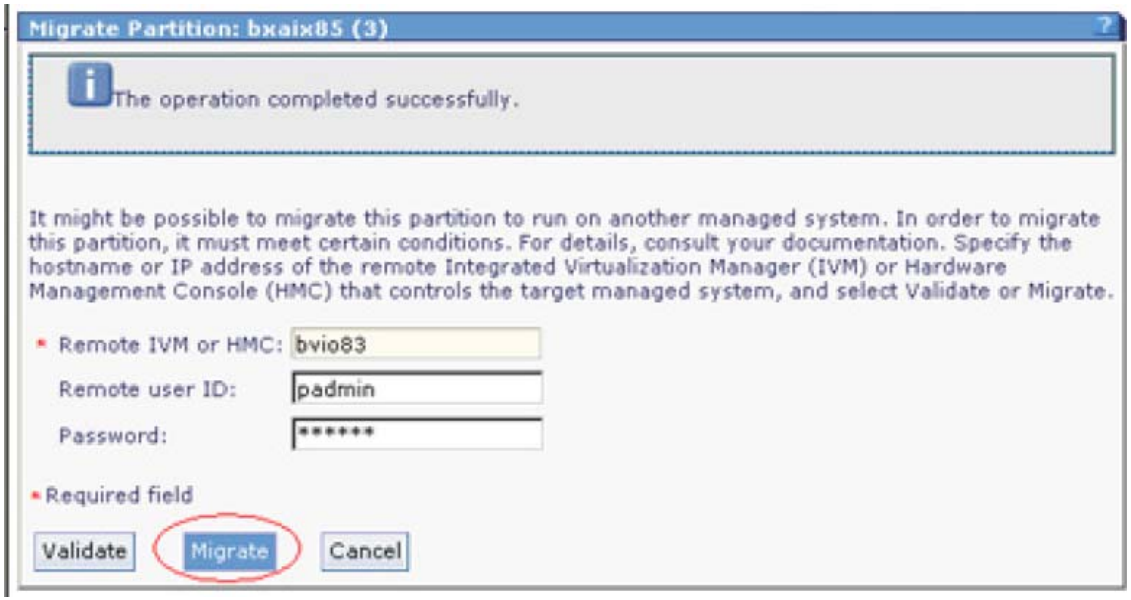
- Presented with a screen to enter the target system details.
- Enter the details and then click on Validate.



- During the validation phase, several configuration checks are performed. Some of the checks include:
  - Ensuring the target system has sufficient memory and processor resources to meet the LPAR's current entitlements.
  - Checking there are no dedicated physical adapters assigned to the LPAR.
  - Verifying that the LPAR does not have any virtual SCSI disks defined as logical volumes on any VIOS. All virtual SCSI disks must be mapped to whole LUNs on the SAN.
  - RMC connections to the LPAR and the source and target VIOS are established.
  - The partition state is active i.e. *Running*. The LPAR's name is not already in use on the target system.
  - A virtual adapter map is generated that maps the source virtual adapter/devices on to the target VIOS. Used during the actual migration.

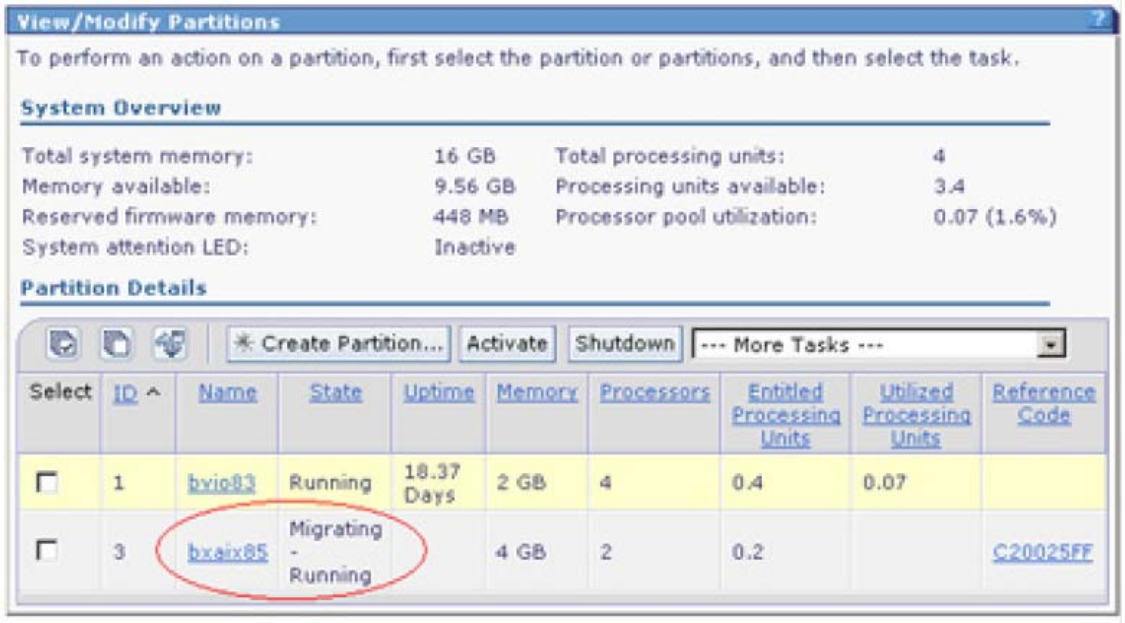
Performing Live Partition Mobility - cont.

- Validation completes successfully.
- Message stating it "*might be possible*“ to migrate the LPAR.
- Click Migrate and the migration to the other Blade begins.



Performing Live Partition Mobility - cont.

- New LPAR created on the target Blade
- Same name as the LPAR on the source Blade.
- State of *Migrating – Running*.



View/Modify Partitions

To perform an action on a partition, first select the partition or partitions, and then select the task.

System Overview

Total system memory:	16 GB	Total processing units:	4
Memory available:	9.56 GB	Processing units available:	3.4
Reserved firmware memory:	448 MB	Processor pool utilization:	0.07 (1.6%)
System attention LED:	Inactive		

Partition Details

Create Partition... | Activate | Shutdown | --- More Tasks ---

Select	ID	Name	State	Uptime	Memory	Processors	Entitled Processing Units	Utilized Processing Units	Reference Code
☐	1	bvio83	Running	18.37 Days	2 GB	4	0.4	0.07	
☐	3	bxaix85	Migrating - Running		4 GB	2	0.2		C20025FF

Performing Live Partition Mobility - cont.

- What happens during the partition migration phase?
  - State information is transferred from the source to the target system.
  - This "state information" includes such things as partition memory, processor state, virtual adapter state, NVRAM.

- Just *some* of the events and actions that occur during the migration:
  - A partition shell is created on the target system.
  - Shell partition reserves resources required to create the inbound LPAR i.e CPU, memory, virtual adapters.
  - A connection between the source and target systems and their respective Hypervisor is established through a device called the Virtual Asynchronous Service Interface (VASI) on the VIOS.
  - The source and target VIOS use this virtual device to communicate with the Hypervisor to gain access to the LPAR's state and to coordinate the migration.
  - You can confirm the existence of this device with the lsdev command on the VIOS.

```
$ lsdev -type adapter | grep vasi
vasi0          Available  Virtual Asynchronous Services Interface (VASI)
```

Performing Live Partition Mobility - cont.

- What happens during the partition migration phase - cont?
  - The *vasistat* command displays the statistics for the VASI device.
  - Run this command on the source VIOS during the migration.
  - Observe "*Total Bytes to Transfer*" indicates the size of the memory copy and that "*Bytes Left to Transfer*" indicates how far the transfer has progressed.

```
$ vasistat -all vasi0 | more
 -------------------------------------------------------------
VASI STATISTICS (vasi0) :
Device Type: Virtual Asynchronous Services Interface (VASI)
Elapsed Time: 21 days 22 hours 2 minutes 7 seconds


Transmit Statistics:                          Receive Statistics:
--------------------                          -------------------
Packets: 11508441                             Packets: 22427345
Bytes: 63442033217                            Bytes: 58347415033
Transmit Errors: 0                            Receive Errors: 0
Bad Packets: 0                                Bad Packets: 0
No Buffers: 0                                 No Buffers: 0
Interrupts: 11508441                          Interrupts: 25020140
                                              System Buffers: 0

Interrupt Processing Exceeded: 253841
Offlevel Interrupt Scheduled: 242986

Driver Flags: Up Running 64BitSupport

Maximum Operations: 8
Maximum Receive Pools: 3
Active Operations: 1

Statistics for each operation:
==============================

Operation #0 (ACTIVE):
----------------------
        Operation Type: Migration (Source)
        Stream ID: DDD1EED073EF6BB1
        TOP/BOTTOM: 00000000/00070000
        Elapsed Time: 0 days 0 hours 0 minutes 1 seconds
        Flags: <RUNNABLE>
        Operation State: Operational
        Stream State: Enabled
        Total Bytes to Transfer: 4364697600
        Bytes Left to Transfer: 4348395520
```

Performing Live Partition Mobility - cont.

- What happens during the partition migration phase - cont?
  - Virtual target devices and virtual SCSI adapters are created on the target system.
  - Using the *lsmap* command on the *target* VIOS, before the migration.
  - Notice that there are no virtual SCSI or virtual target device mappings.

```
login as: padmin
padmin@bvio83's password:
Last unsuccessful login: Fri Aug 15 15:19:24 EET 2008 on ssh from 10.0.118.181
Last login: Tue Aug 26 21:24:50 EET 2008 on ssh from 10.0.9.201
$ lsmap -all
$
```

Performing Live Partition Mobility - cont.

- What happens during the partition migration phase - cont?
  - Running the same command after the migration shows that the virtual disk mappings have been automatically created, as part of the migration process.

```
login as: padmin
padmin@bvio83's password:
Last unsuccessful login: Fri Aug 15 15:19:24 EET 2008 on ssh from 10.0.118.181
Last login: Tue Aug 26 21:21:23 EET 2008 on ssh from 10.3.29.156
$ lsmap -all
SVSA            Physloc                                     Client Partition ID
--------------- ------------------------------------------- ------------------
vhost0          U7998.61X.10071FA-V1-C11                    0x00000002

VTD                   vtscsi0
Status                Available
LUN                   0x8200000000000000
Backing device        hdisk1
Physloc               U78A5.001.WIH074C-P1-C6-T1-W500507630603059A-L4000402D0000
0000

VTD                   vtscsi1
Status                Available
LUN                   0x8300000000000000
Backing device        hdisk2
Physloc               U78A5.001.WIH074C-P1-C6-T1-W500507630603059A-L4001402D0000
0000

VTD                   vtscsi2
Status                Available
LUN                   0x8100000000000000
Backing device        hdisk3
Physloc               U78A5.001.WIH074C-P1-C6-T1-W500507630603059A-L4000402E0000
0000
```

Performing Live Partition Mobility - cont.

- What happens during the partition migration phase - cont?
  - The LPAR's physical memory pages are copied to the *shell* LPAR on the target system.
  - Using the *topas* command on the *source* VIOS, observe network traffic on the SEA (ent6) as a result of the memory copy.

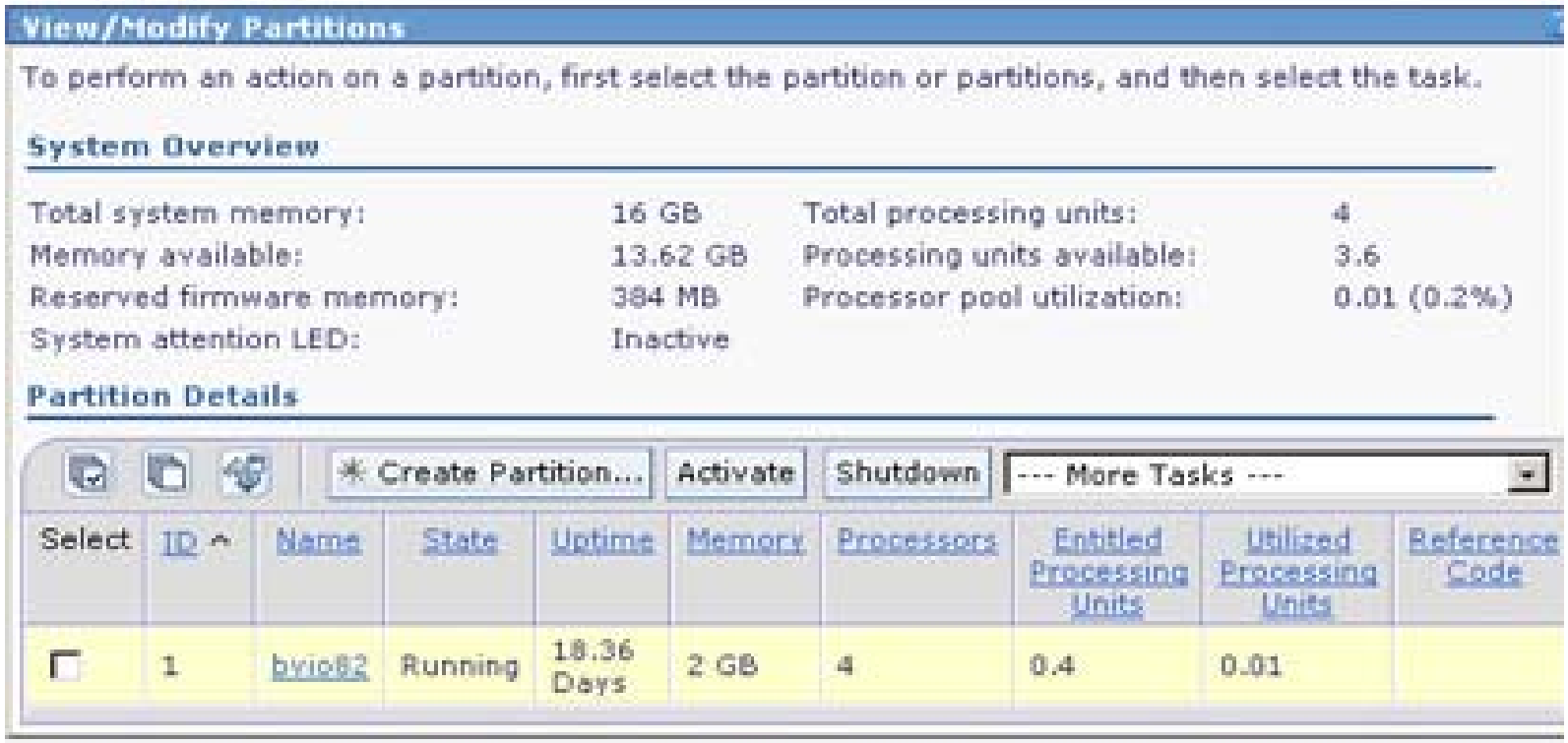
```
Topas Monitor for host:    bvio82                EVENTS/QUEUES    FILE/TTY
Thu Sep 18 13:47:36 2008   Interval:  2          Cswitch   26317  Readch     1148
                                                 Syscall    2362  Writech    6532
Kernel   90.0    |##########################  |  Reads        19  Rawin         0
User      0.6    |#                           |  Writes       40  Ttyout      223
Wait      0.0    |                            |  Forks         0  Igets         0
Idle      9.4    |###                         |  Execs         0  Namei        81
Physc =  1.36                   %Entc= 339.8    Runqueue     1.5  Dirblk        0
                                                 Waitqueue    0.0
Network  KBPS    I-Pack  O-Pack   KB-In  KB-Out
en6      80.4K     32.6K   64.0K 3433.1    77.1K PAGING           MEMORY
lo0       0.0       0.0     0.0     0.0     0.0  Faults       112  Real,MB    2048
                                                 Steals         0  % Comp     52.2
Disk    Busy%     KBPS      TPS KB-Read KB-Writ  PgspIn         0  % Noncomp  40.0
hdisk2    0.0      8.0      0.5     0.0     8.0  PgspOut        0  % Client   40.0
hdisk0    0.0      0.0      0.0     0.0     0.0  PageIn         0
                                                 PageOut        0  PAGING SPACE
Name            PID  CPU%  PgSp Owner           Sios           0  Size,MB    1536
ctrlproc   31887534  29.6   0.9 root                              % Used      9.9
seaproc      131174  28.1   1.0 root            NFS (calls/sec)  % Free     91.1
accesspr     282772   7.5   1.1 root            ServerV2       0
java         286882   3.6 101.9 root            ClientV2       0   Press:
topas      31940832   3.3  51.6 root            ServerV3       0   "h" for help
syslogd      393422   1.0   0.2 root            ClientV3       0   "q" to quit
```

Performing Live Partition Mobility - cont.

- What happens during the partition migration phase - cont?
  - LPAR is still active. SAP still running. State continues to change while the memory is copied.
  - Memory pages modified during the transfer marked as *dirty*.
  - Process is repeated until dirty pages no longer decreasing.
  - At this point *target* system instructs Hypervisor on *source* to suspend LPAR.
  - During suspension, *source* LPAR continues to send state information to *target*.
  - LPAR is then *resumed*.
  - Resumes execution on the *target* system.
  - If LPAR requires a page not yet been migrated, will be "demand-paged“ from the *source* system.
  - The LPAR recovers its I/O operations.
  - A gratuitous ARP request is sent on all virtual Ethernet adapters to update the ARP caches on all external switches and systems on the network.
  - The LPAR is now active again.
  - Target system receives the last dirty page from the source system, the migration is complete.
  - Period between suspension and resumption lasts just a *few seconds*.
  - During my tests, I did not notice *any* disruption to the LPAR as a result of this operation.

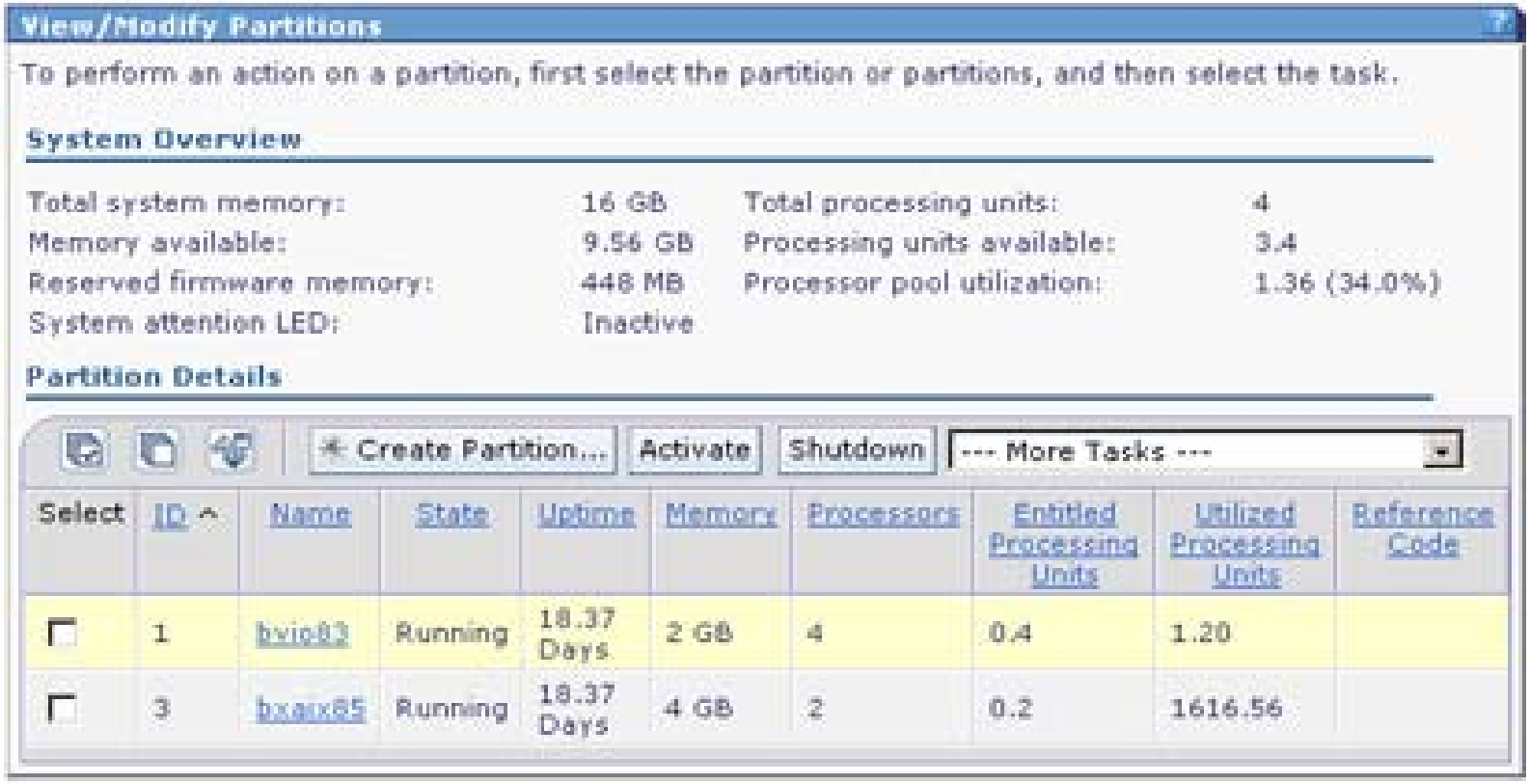
Performing Live Partition Mobility - cont.

- Memory copy complete. VIOS on *source* removes virtual SCSI server adapters and device to LUN mapping that existed previously.
- LPAR is automatically deleted from the *source* Blade.



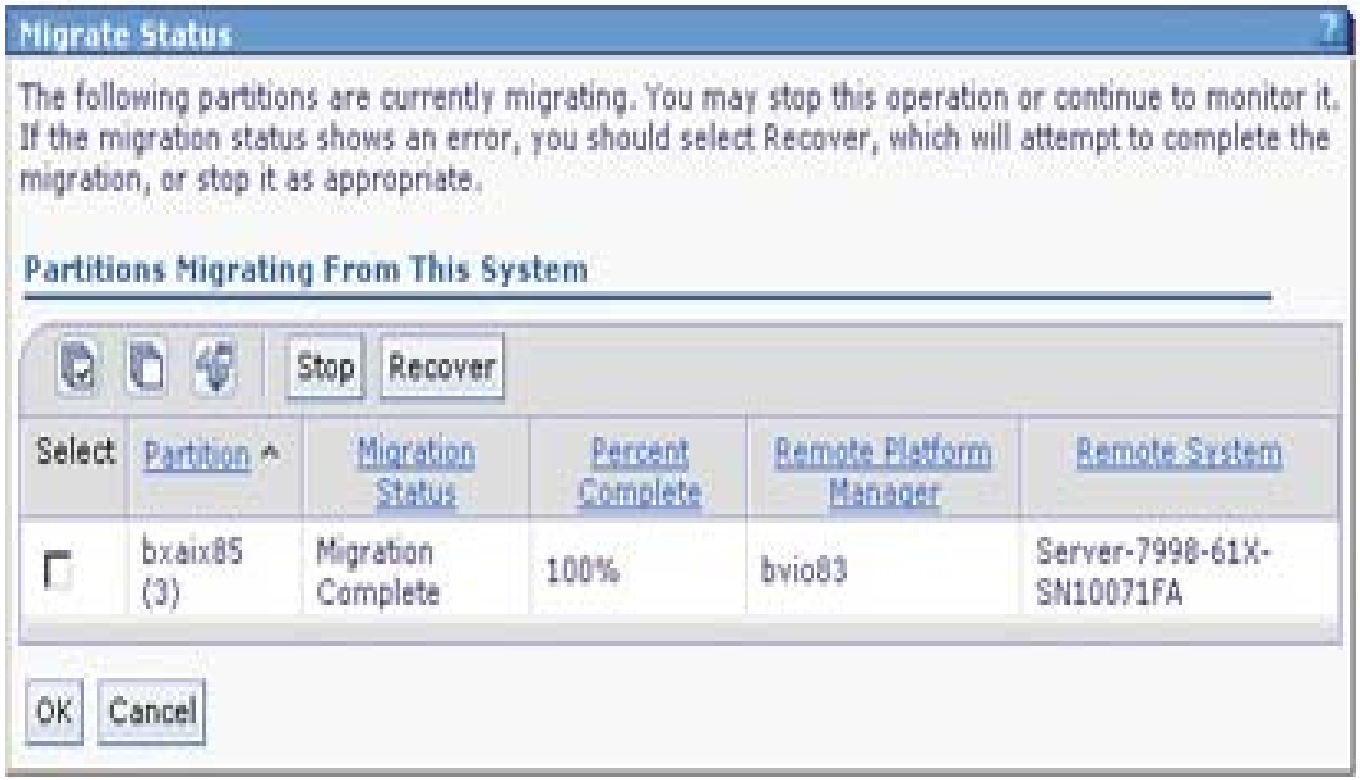
Performing Live Partition Mobility - cont.

- LPAR is now in a *Running* state on the *target* Blade.



Performing Live Partition Mobility - cont.

- The migration is 100% complete.



Verifying Live Partition Mobility.

- Now that the LPAR is running on the other Blade, run the *lsconf* command again to confirm that the serial number has changed with the physical hardware:

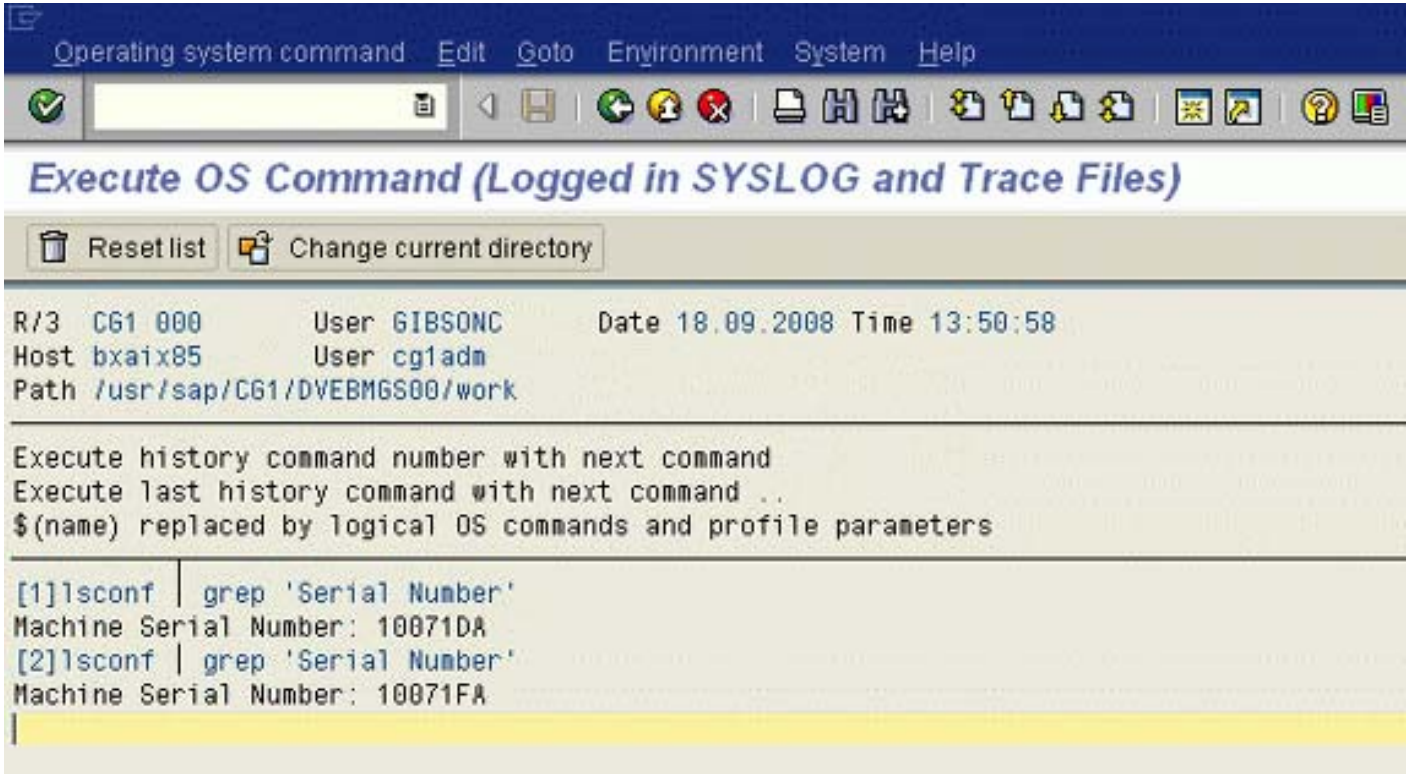
```
gibsonc@bxaix85 /home/gibsonc $ lsconf | grep 'Serial Number'
Machine Serial Number: 10071FA
```

- In order to confirm and verify that SAP and Oracle are not impacted by the migration, check the Oracle alert log for any errors. No errors are found.

```
root@bxaix85 / # tail -f /oracle/CG1/saptrace/background/alert_CG1.log
  Current log# 4 seq# 1396 mem# 1: /oracle/CG1/mirrlogB/log_g14m2.dbf
Thu Sep 18 13:45:52 2008
Completed checkpoint up to RBA [0x572.2.10], SCN: 9878246
Thu Sep 18 13:47:15 2008
Beginning log switch checkpoint up to RBA [0x575.2.10], SCN: 9907535
Thread 1 advanced to log sequence 1397
  Current log# 1 seq# 1397 mem# 0: /oracle/CG1/origlogA/log_g11m1.dbf
  Current log# 1 seq# 1397 mem# 1: /oracle/CG1/mirrlogA/log_g11m2.dbf
Thu Sep 18 13:48:31 2008
Completed checkpoint up to RBA [0x573.2.10], SCN: 9888598
```

Verifying Live Partition Mobility - cont.

- From within SAP, run the *lsconf* command before and after the migration to confirm that the physical server has changed:



Post Migration.

- My ssh login sessions on bxaix85 remained active.
- SAP team did not notice any disruption to their SAP GUI client sessions or jobs running on the LPAR.
- Mobility activity is logged on the LPAR and the source and target VIOS.
- Review the logs with the errpt (AIX) and errlog (VIOS) commands.
- On AIX you'll notice messages similar to CLIENT_PMIG_STARTED and CLIENT_PMIG_DONE.
- Additional information from DRMGR, on AIX, is also logged to syslog,
- For instance, Starting CHECK phase for partition migration.
- On the VIOS you'll find messages relating to the suspension of the LPAR and the migration status (Client partition suspend issued and Migration completed successfully).
- The final objective has been achieved. LPAR running on a different physical server. Can now perform scheduled maintenance activities on the Blade.
- SAP will not suffer any down time as a result of this activity.

LPM POC complete.

- POC successful. Migration took roughly two minutes to complete.
- LPAR being moved had 4GB of memory. Time required for copying of the LPAR's memory from the source to the target.
- The "suspend" of the LPAR itself lasted no more than two seconds.
- Considerations:
  - Using a high-performance network between the source and target systems.
  - Prior to migration, recommend reducing the LPAR's memory update activity.
  - These steps will help to improve the overall performance of the migration.
  - We used a 1GB network within our Blade environment.
  - For larger System p servers (570 and 595), we are considering using a 10GB network when we start moving systems with a large amount of memory (80GB or more). Are we likely to do this?
- LPM enormous potential for reducing scheduled downtime for maintenance activities.
- No disruption to user applications.
- Power to adjust resource usage. LPARs can moved to different servers to balance workload.
- Migration/consolidation (POWER6 to POWER7) easier. Simply move the LPAR to POWER7.

Verifying LPM on the 570 and 595s

- Procedures defined from our POC. Verify them with 570 and 595.
- Prior to executing a LPM request, the following requirements must be met:
  - LPAR must be a non-production system.
  - LPAR must be fully virtualised. No physical devices.
  - Virtual disk devices must be connected via the SAN.
  - The VIOC storage must be zoned appropriately on all participating VIOS.
  - UNIX health check scripts must be disabled in root's crontab. They should be enabled again after the migration.
  - VIOS VLAN ids should match on both the source and destination VIOS.
  - Disable virtual SCSI disk health checks prior to the migration. Re-enable the check after the migration.
  - Sufficient spare CPU capacity to cater for the LPAR. 'lparstat –i'
  - Sufficient spare Memory capacity to cater for the LPAR. 'lparstat –i'
  - Recommended LPM performed during "quiet" period i.e. off peak workload.
  - An approved Remedy and/or Change Request.

Verifying LPM on the 570 and 595s - cont

- Disable UNIX health check scripts in root's crontab:

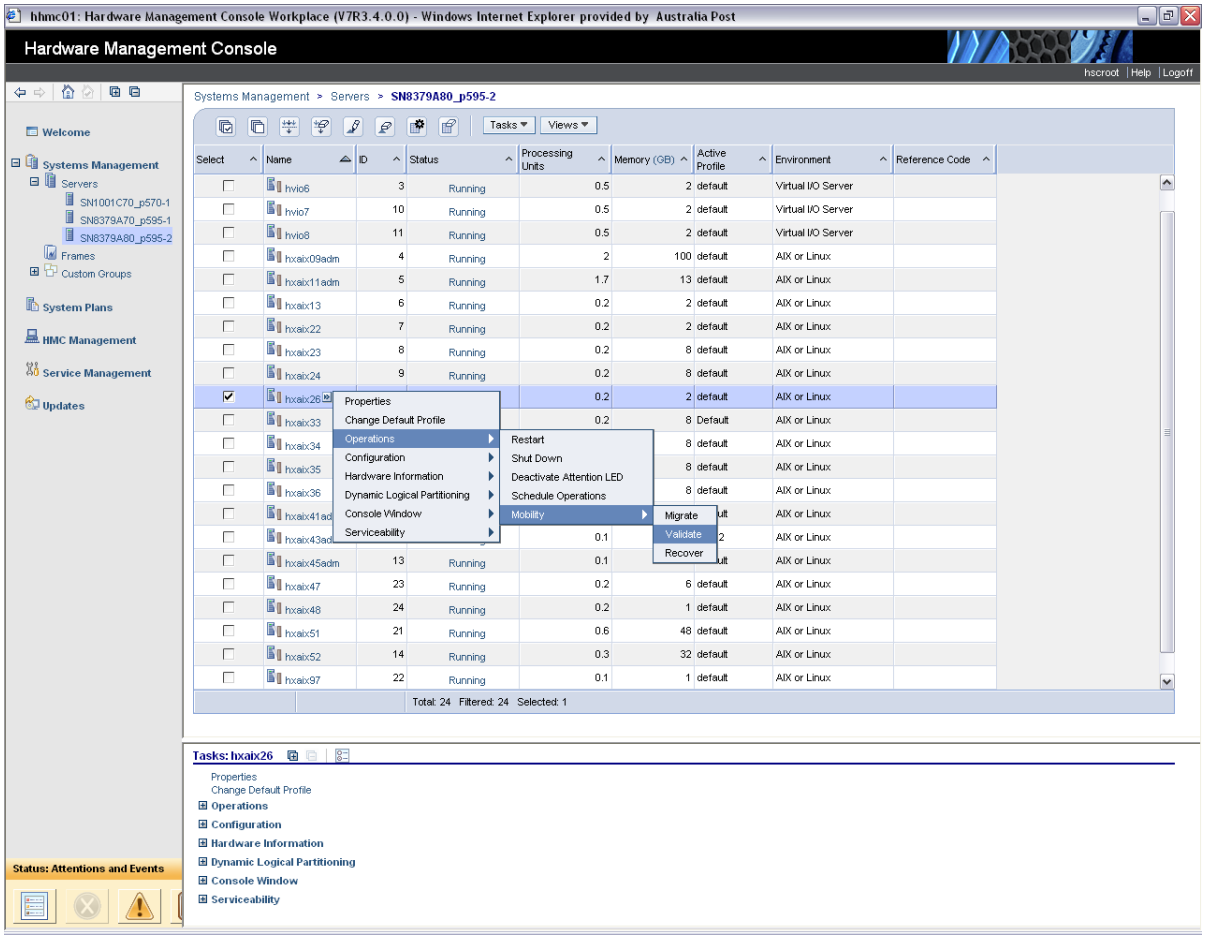
```
#00 19 * * * /usr/local/bin/saveskelvg >> /var/log/saveskelvg.log 2>&1
#00 01 * * 0,3 /usr/local/bin/mksysb2nim >> /var/log/mksysb2nim.log 2>&1
#00 05 * * 0 /usr/local/bin/AIXinfo -repos >> /var/log/chksys.log 2>&1
#00 07 * * * /home/nmon/nmon.ksh > /dev/null 2>&1
#0 16 * * 0 /usr/local/adm/backup.ksh -online -unix -archive 1>/dev/null 2>&1
#0,30 * * * * /usr/local/adm/health_check -check 1>/dev/null 2>&1
```

- Disable virtual SCSI disk health checks prior to the migration:

```
# chdev -l hdiskX -a hcheck_interval=0 -P
```

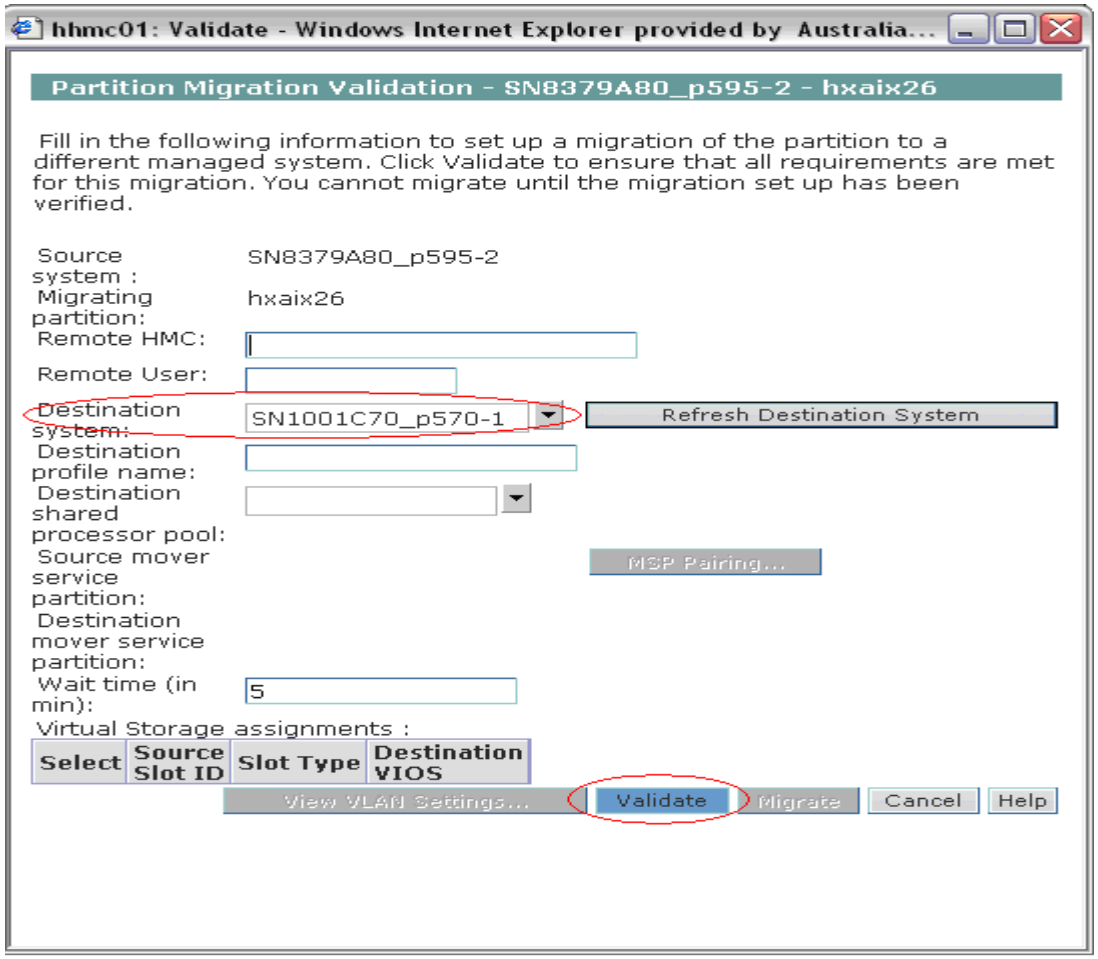
Performing LPM on the 570 and 595s

- Move LPAR from 595-2 to 570-1.
- Validation via HMC similar to IVM. Select the LPAR to migrate e.g. hxaix26.



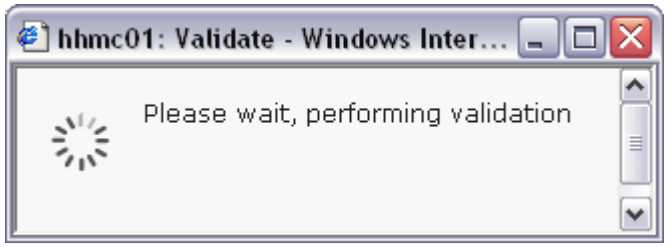
Performing LPM on 570 and 595s – cont.

- Confirm the destination system is correct e.g. SN1001C70_p570-1 is 570-1. Select Validate.

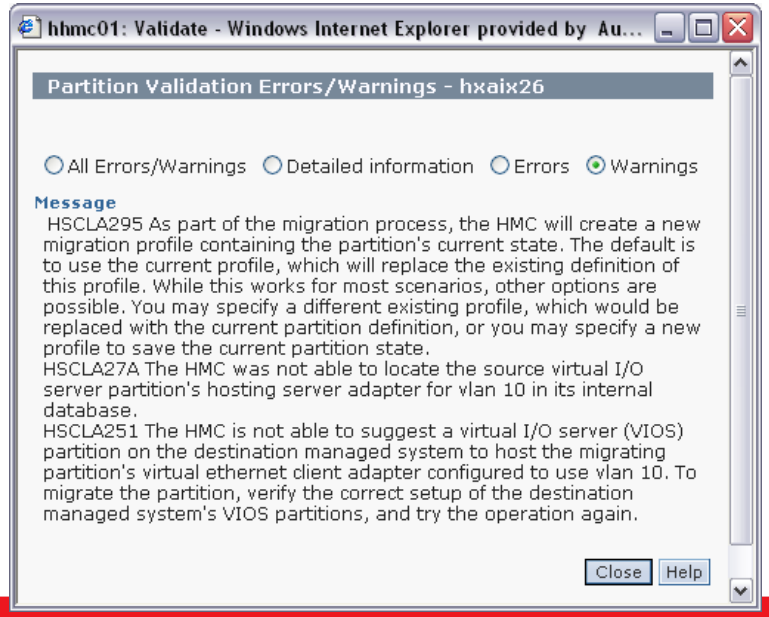


LPM Validation – continued.

- The validation may take several minutes.

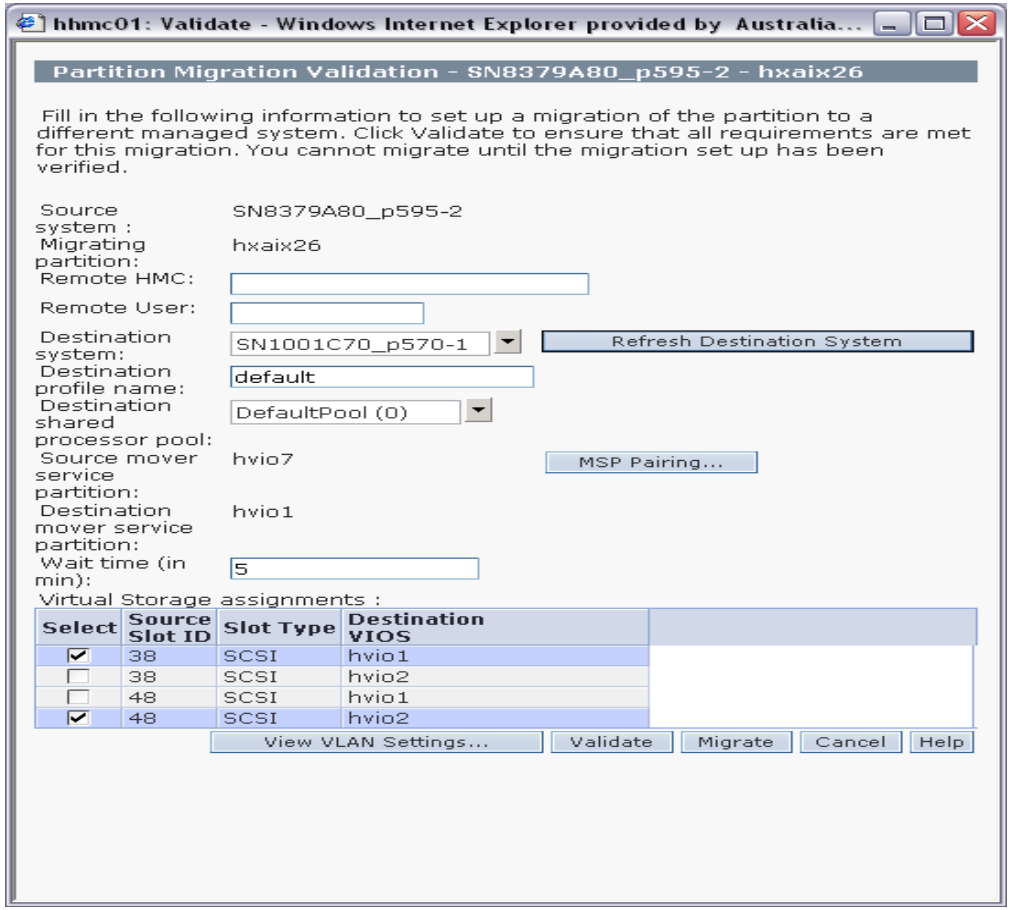


- Ignore messages relating to creating a profile and *vlan 10*. Other messages should be investigated.



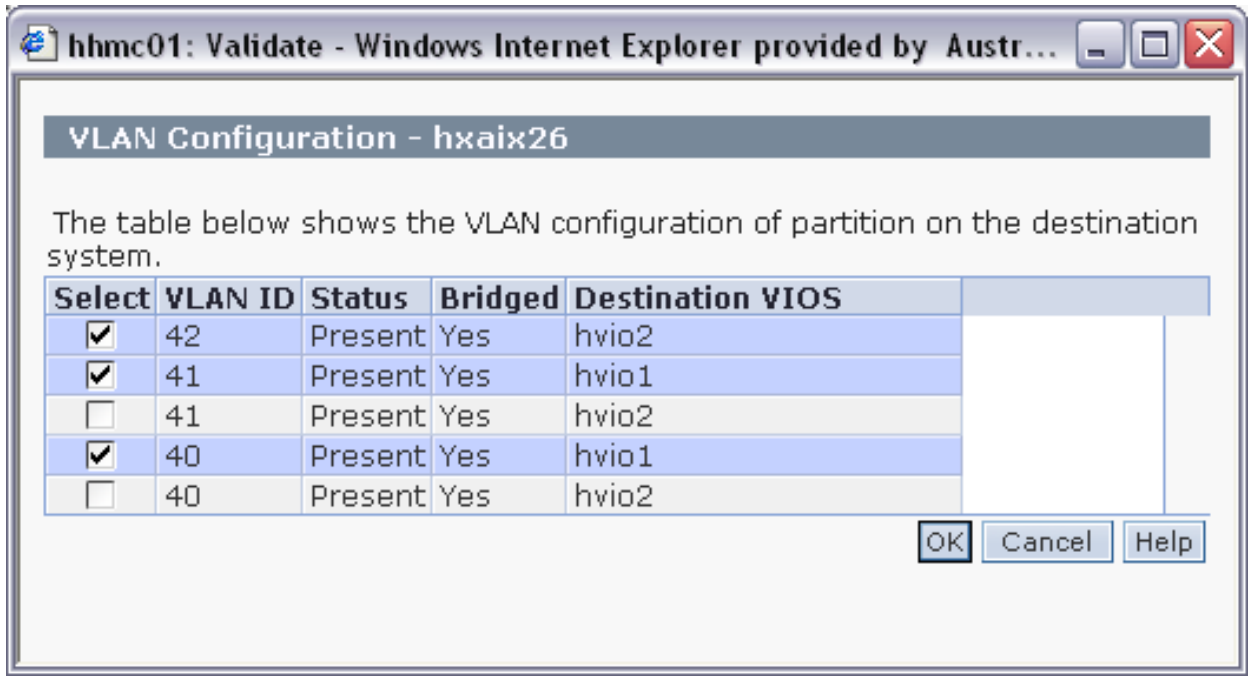
LPM Validation – continued.

- Verify that the appropriate VIOS and virtual storage slots have been selected e.g. Destination VIOS hvio1, slot 38 and hvio2, slot 48.



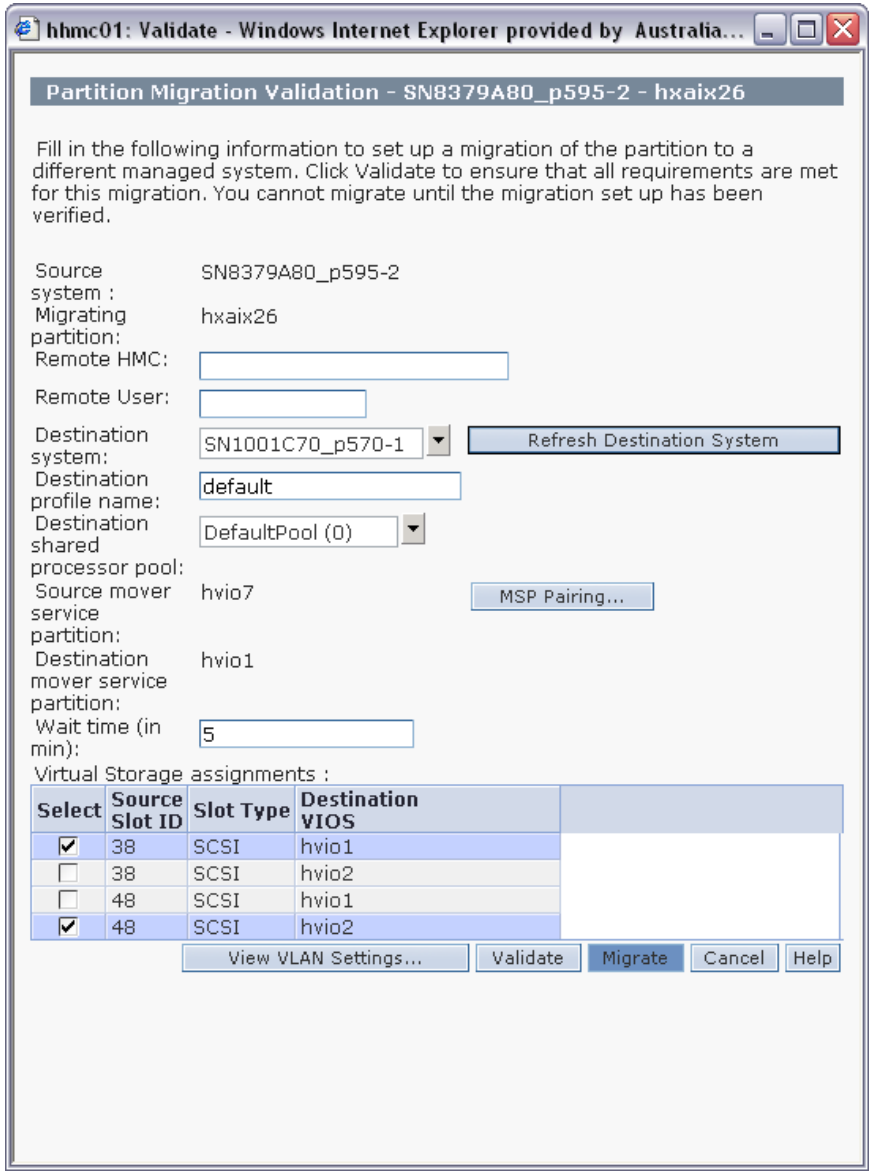
LPM Validation – continued.

- Verify that the correct virtual I/O VLANs have been selected e.g. destination VIOS hvio1, vlan 40 & 41 and hvio2, vlan 42.



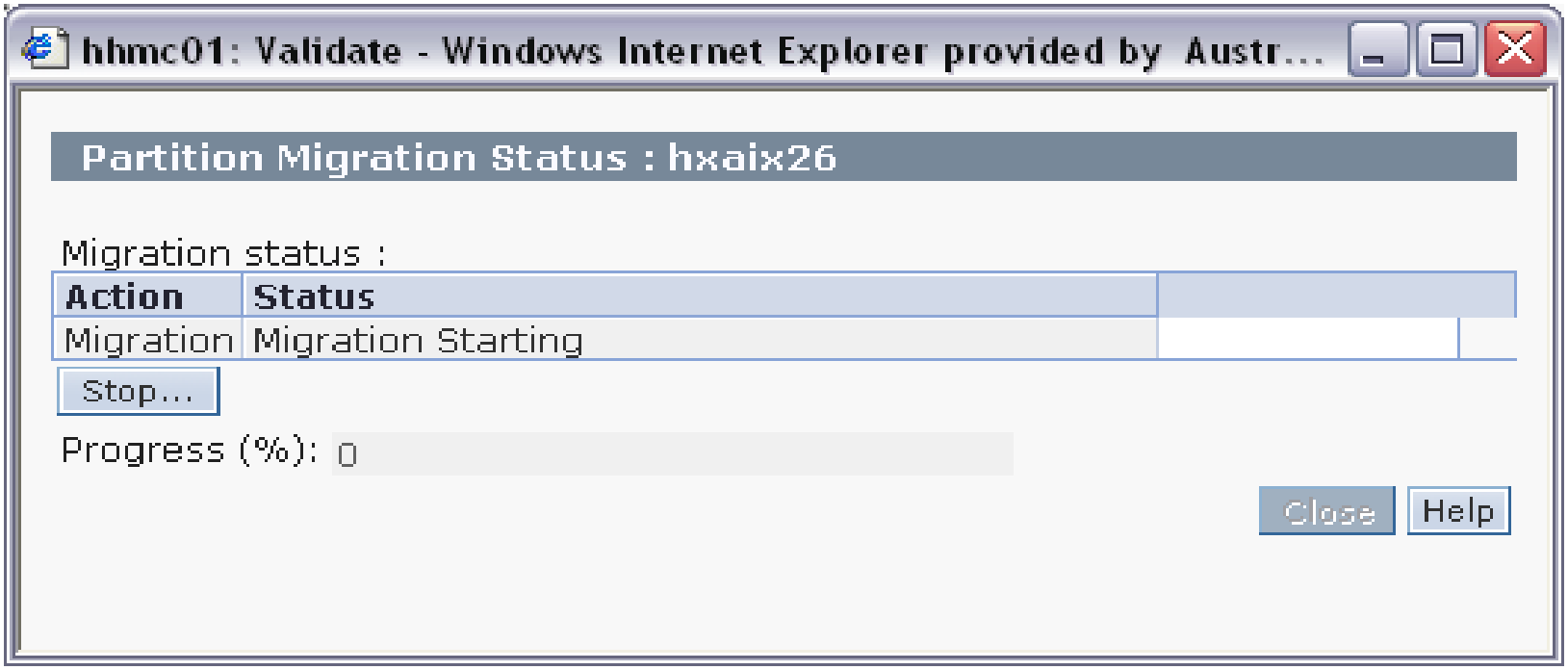
LPM – Migration.

- Execute the Migration.
- Select *Migrate*.



LPM – Migration – continued.

- The migration may take several minutes e.g. 5 minutes or more.

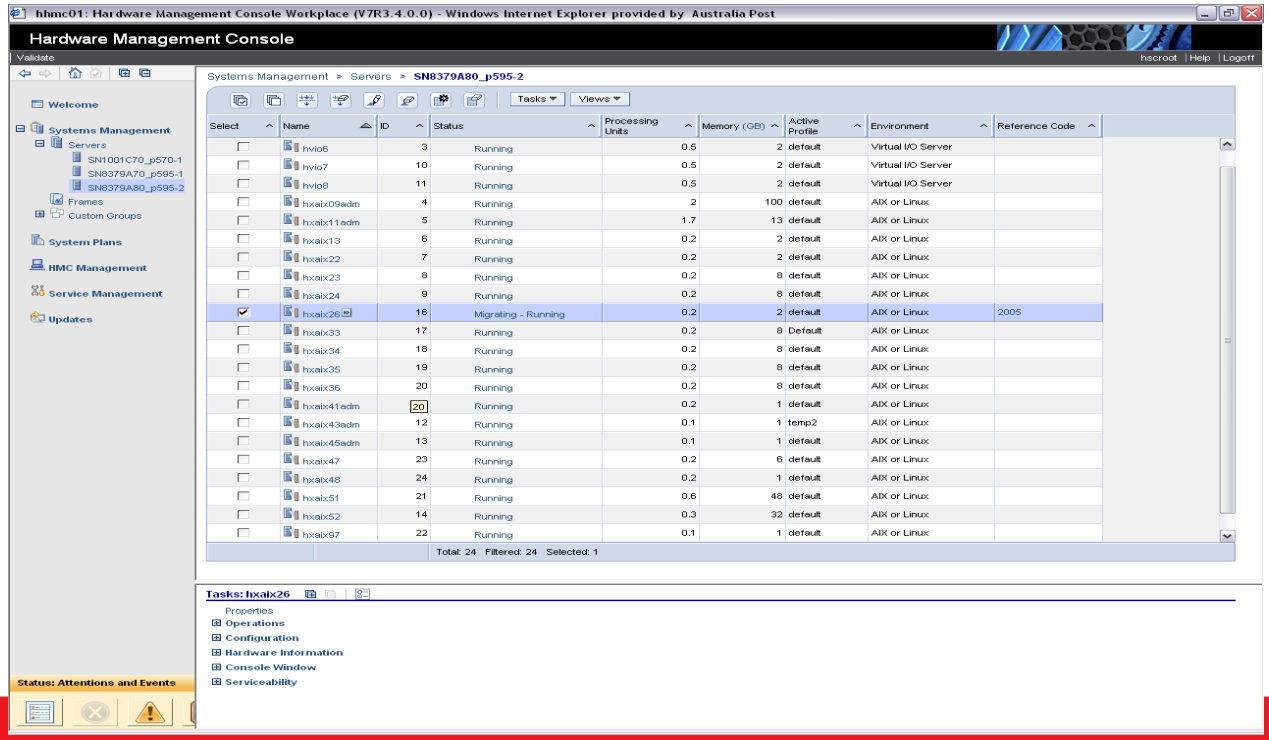


LPM – Migration – continued.

- View the current (source) systems serial number. This will change after the migration.

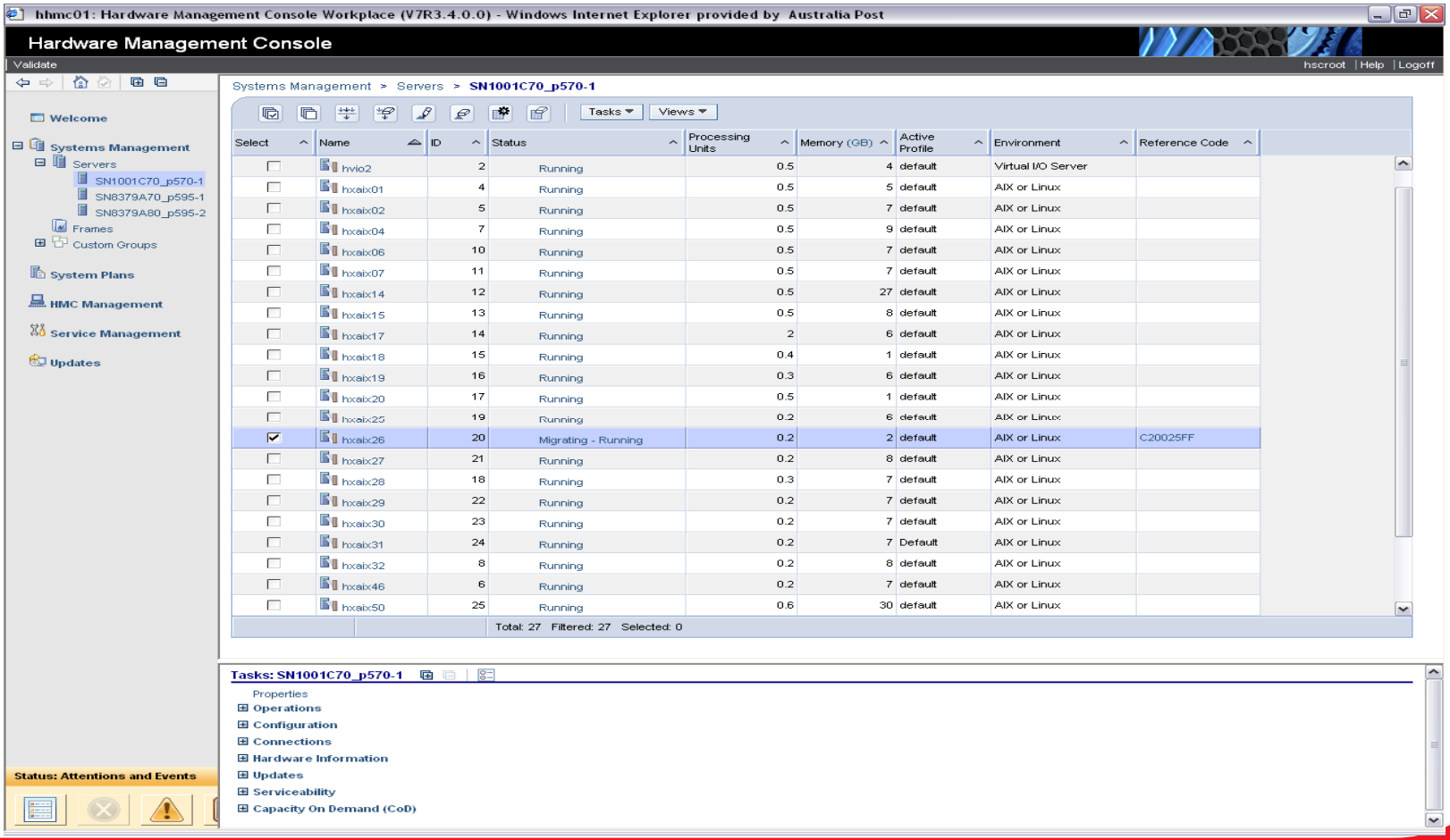
```
root@hxaix26 / # lsattr -El sys0 -a systemid
systemid IBM,028379A80 Hardware system identifier False
```

- The state of the LPAR, hxaix26, on the source system (595-2) will show ‘Migrating –Running’.



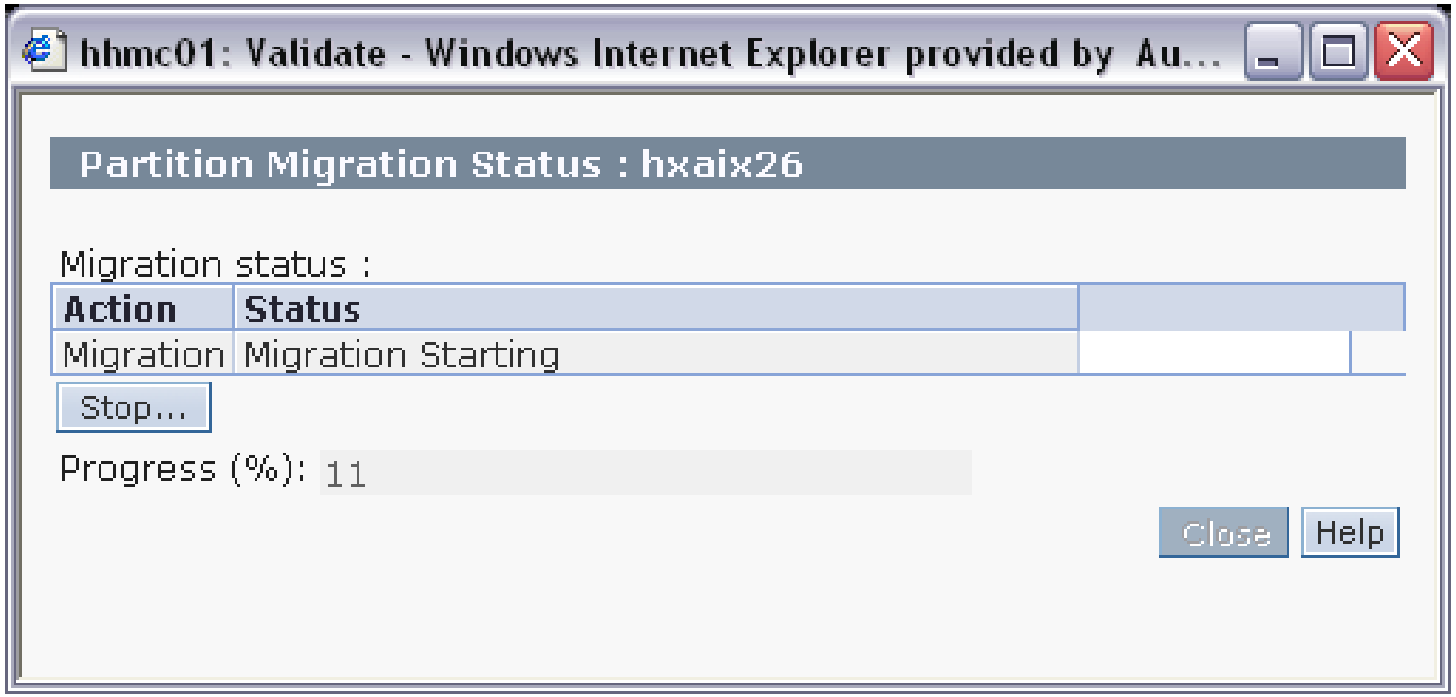
LPM – Migration – continued.

- The state of the *shell* LPAR, hxaix26, on the destination system (570-1) will also show '*Migrating –Running*'.



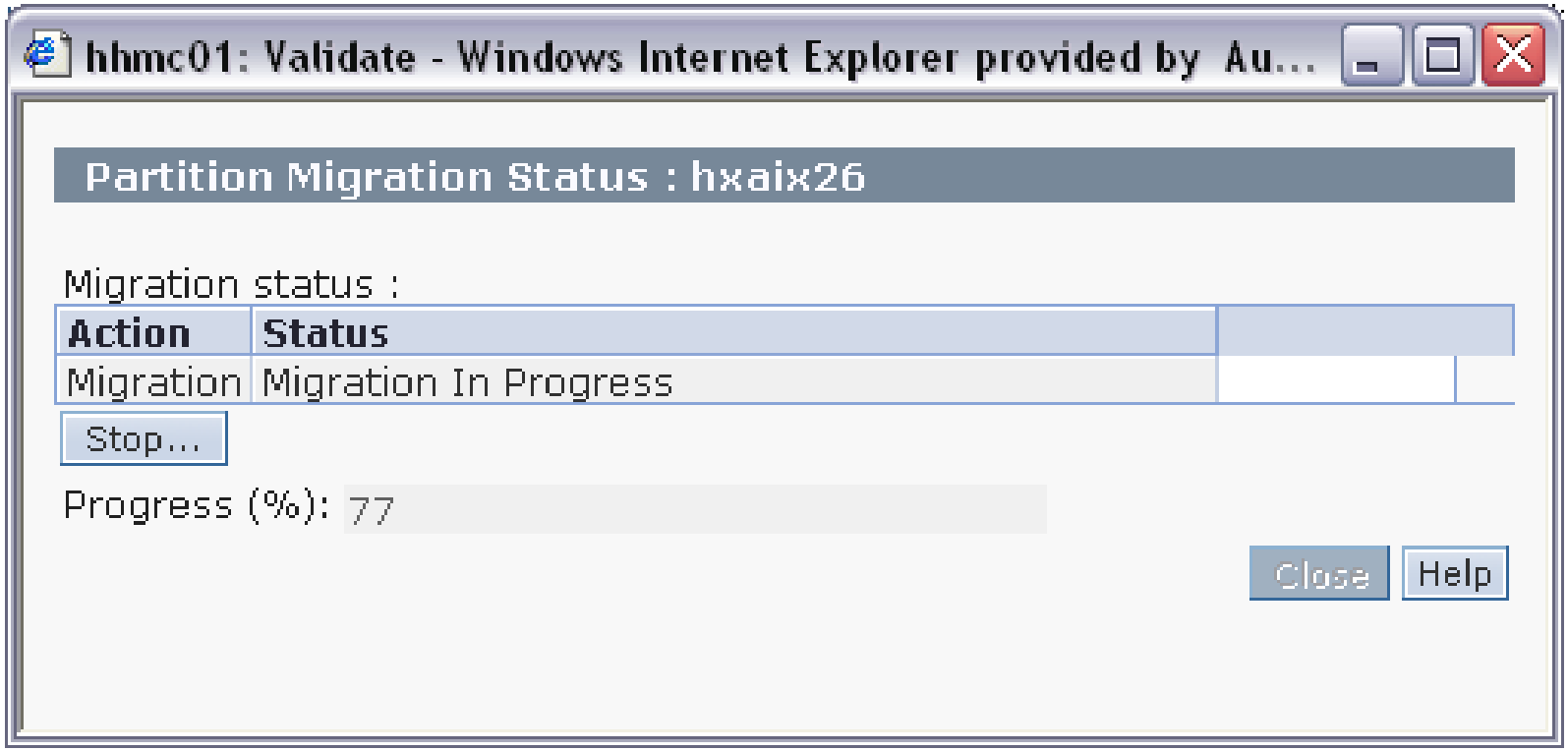
LPM – Migration – continued.

- Monitor the progress and status of the migration.



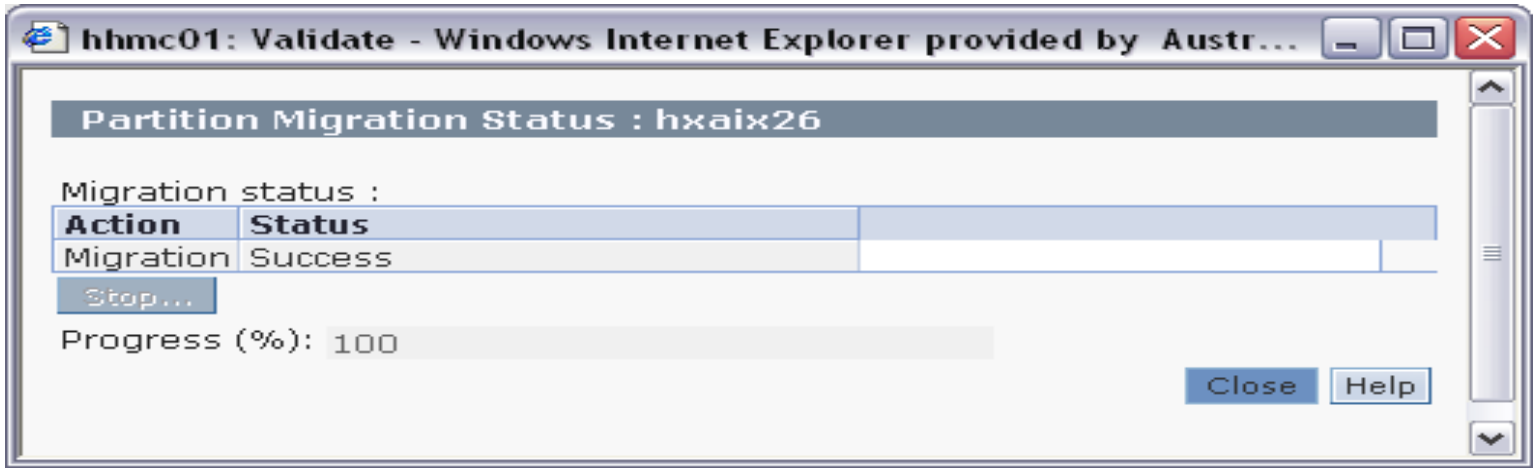
LPM – Migration – continued.

- When the status shows ‘Migration In Progress’, the LPAR will be active on the destination system momentarily. A slight pause may be observed on the LPAR.



LPM – Migration complete.

- The migration is complete, once the status changes to ‘Success’. Click ‘Close’.



Verification

- Confirm that the LPAR has moved to the destination system. The systems serial number should have changed accordingly.

```
root@hxaix26 / # lsattr -El sys0 -a systemid
systemid IBM,021001C70 Hardware system identifier False
```

- The LPAR is removed from the source system (595-2). Now in “*Running*” state on the destination system (570-1).

Post Migration

- The following errors may appear in the AIX error report. Messages relating to path failures and the system planar are expected.

```
root@hxaix26 / # errpt
IDENTIFIER TIMESTAMP  T C RESOURCE_NAME  DESCRIPTION
DE3B8540   0210141209 P H hdisk10        PATH HAS FAILED
DE3B8540   0210141209 P H hdisk7         PATH HAS FAILED
DE3B8540   0210141209 P H hdisk9         PATH HAS FAILED
DE3B8540   0210141209 P H hdisk8         PATH HAS FAILED
DE3B8540   0210141209 P H hdisk3         PATH HAS FAILED
DE3B8540   0210141209 P H hdisk2         PATH HAS FAILED
DE3B8540   0210141209 P H hdisk6         PATH HAS FAILED
DE3B8540   0210141209 P H hdisk1         PATH HAS FAILED
DE3B8540   0210141209 P H hdisk5         PATH HAS FAILED
DE3B8540   0210141209 P H hdisk4         PATH HAS FAILED
A5E6DB96   0210141209 I S pmig           Client Partition Migration Completed
CC40E579   0210141109 P H sysplanar0     ELECTRICAL POWER RESUMED
```

Post Migration – continued.

- The following messages will appear in syslog.

```
root@hxaix26 / # tail -f /var/log/syslog
Feb 10 13:59:26 hxaix26 user:info syslog:  ~~~~ Start: DR PMIG operation ~~~~
Feb 10 13:59:26 hxaix26 user:info syslog: Starting partition migration PRE phase
Feb 10 14:11:37 hxaix26 local1:info DRMGR: Starting phase PRE for kernel.
Feb 10 14:11:37 hxaix26 local1:info DRMGR: Starting PRE phase for scripts.
Feb 10 14:11:39 hxaix26 local1:info DRMGR: Completed the phase for Scripts.
Feb 10 14:11:39 hxaix26 local1:info DRMGR: Starting phase PRE for signal delivery.
Feb 10 14:11:40 hxaix26 local1:info DRMGR: Completed signal delivery phase.
Feb 10 14:11:40 hxaix26 local1:info DRMGR: Completed PRE signal phase.
Feb 10 14:11:40 hxaix26 local1:info DRMGR: Starting phase PRE for kernel extensions.
Feb 10 14:11:40 hxaix26 local1:info DRMGR: Completed the phase for kernel extensions.
Feb 10 14:11:40 hxaix26 local1:info DRMGR: Starting Kernel phase.
Feb 10 14:11:40 hxaix26 local1:info DRMGR: Starting phase KERNEL for kernel.
Feb 10 14:11:57 hxaix26 daemon:warn|warning inetd[225476]: Server /usr/bin/xmtopas
  has ended with exit status 0x4100.
Feb 10 14:12:10 hxaix26 local1:info DRMGR: Starting POST phase.
Feb 10 14:12:10 hxaix26 local1:info DRMGR: Starting phase POST for kernel.
Feb 10 14:12:10 hxaix26 local1:info DRMGR: Starting phase POST for kernel extensions.
Feb 10 14:12:10 hxaix26 local1:info DRMGR: Completed the phase for kernel extensions.
Feb 10 14:12:10 hxaix26 local1:info DRMGR: Starting phase POST for signal delivery.
Feb 10 14:12:20 hxaix26 local1:info DRMGR: 1 applications have not handled signals
  yet
Feb 10 14:12:21 hxaix26 local1:info DRMGR: Completed signal delivery phase.
Feb 10 14:12:22 hxaix26 local1:info DRMGR: Starting POST phase for scripts.
Feb 10 14:12:49 hxaix26 local1:info DRMGR: Completed post notification for DLPAR
  scripts.
Feb 10 14:12:49 hxaix26 local1:info DRMGR:  ~~~~ End: DR PMIG operation ~~~~
```

Post Migration – continued.

- Enable UNIX health check scripts in root's crontab.

```
00 19 * * * /usr/local/bin/saveskelvg >> /var/log/saveskelvg.log 2>&1
00 01 * * 0,3 /usr/local/bin/mksysb2nim >> /var/log/mksysb2nim.log 2>&1
00 05 * * 0 /usr/local/bin/AIXinfo -repos >> /var/log/chksys.log 2>&1
00 07 * * * /home/nmon/nmon.ksh > /dev/null 2>&1
0 16 * * 0 /usr/local/adm/backup.ksh -online -unix -archive 1>/dev/null 2>&1
0,30 * * * * /usr/local/adm/health_check -check 1>/dev/null 2>&1
```

- Re-enable virtual SCSI disk health checks after the migration.

```
# chdev -l hdiskX -a hcheck_interval=60 –P
```

- A tip regarding performance tools (i.e. topas) & LPM.
- Can run tools during migration but be aware data reported may not be meaningful.
- Underlying server hardware changes, performance counters are likely to be reset.
- I observed my topas session reset itself when the migration had completed.

```
Partition migration is over.....restarting topas
```

Known Problems.

- Problem
  - Following a migration the ***`lsmap -all`*** command does not show the correct partition ID.
  - If the ID of the Mobile partition changes as part of migration, the connected partition ID is not updated to show the correct number.
  - There is no functional problem caused, rebooting the mobile partition will cause the correct partition ID to be displayed.

- Workaround
  - Create an ID for the partition that is unique across all systems

Known Problems - continued.

- Problem
  - Refresh of IBM.ManagementServer resource failed. Error report shows:
    `DD42A684  I S DRMGR    DR script related Message.`
- Workaround

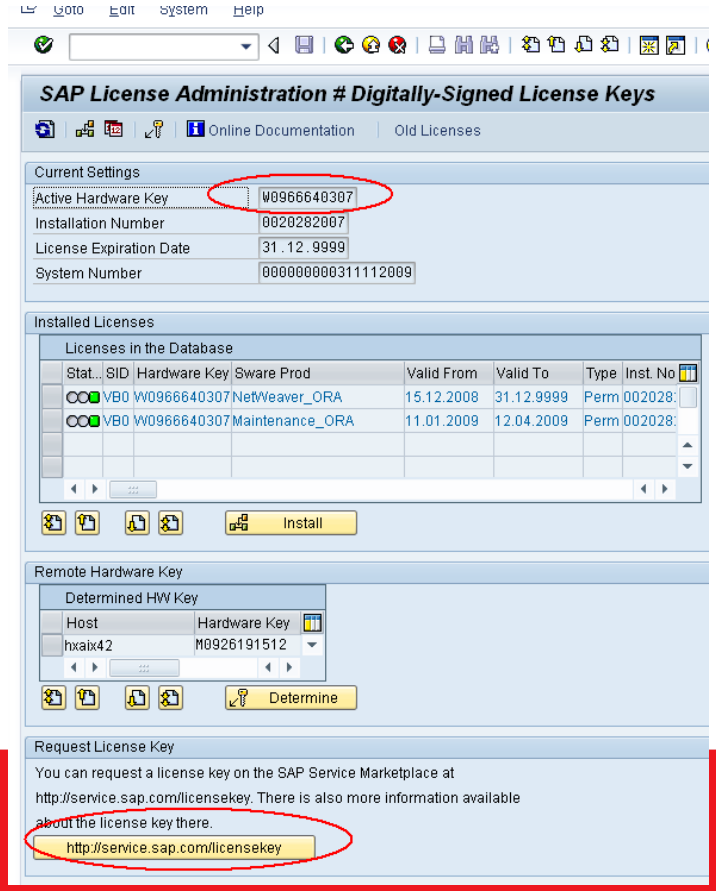
```
# chown root.system /usr/lib/dr/scripts/all/IBM.CSMAgentRM_dr.sh
```

Limitations and Restrictions.

- Virtual target device names
  - Custom names for virtual target devices in VIOS are lost when a partition is migrated. If custom names have been assigned to devices that belong to a partition that is migrated, the migrated virtual target devices will use default names.
- Virtual SCSI Server adapter IDs
  - When a partition is migrated, the VSCSI Server adapters are automatically assigned adapter numbers. Currently it is not possible to specify the slots to be used. The adapters are assigned to use the next available slot.
- Refer to the following IBM website for further information:
  http://www14.software.ibm.com/webapp/set2/sas/f/pm/known.html

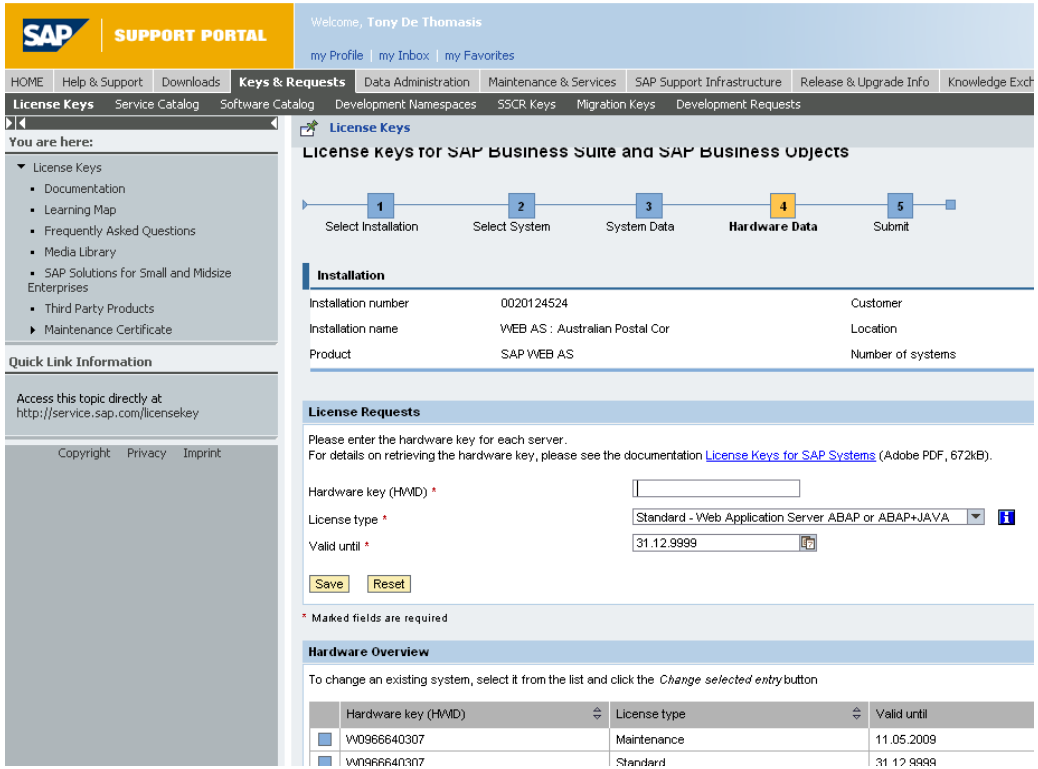
SAP License Keys

- The SAP Instance Hardware key will change after a successful LPM. As long as the SAP instance is not re-started the existing SAP license key will be retained in the SAP memory buffer, therefore no licence key problems are expected.
- Should an additional license key be required, please register the new hardware key (after LPM) via the SAP Service Marketplace http://service.sap.com/licensekey, and then apply the new license key to the SAP instance using the SLICENSE transaction. The whole process should take less than 30 minutes.



SAP License Keys – cont.

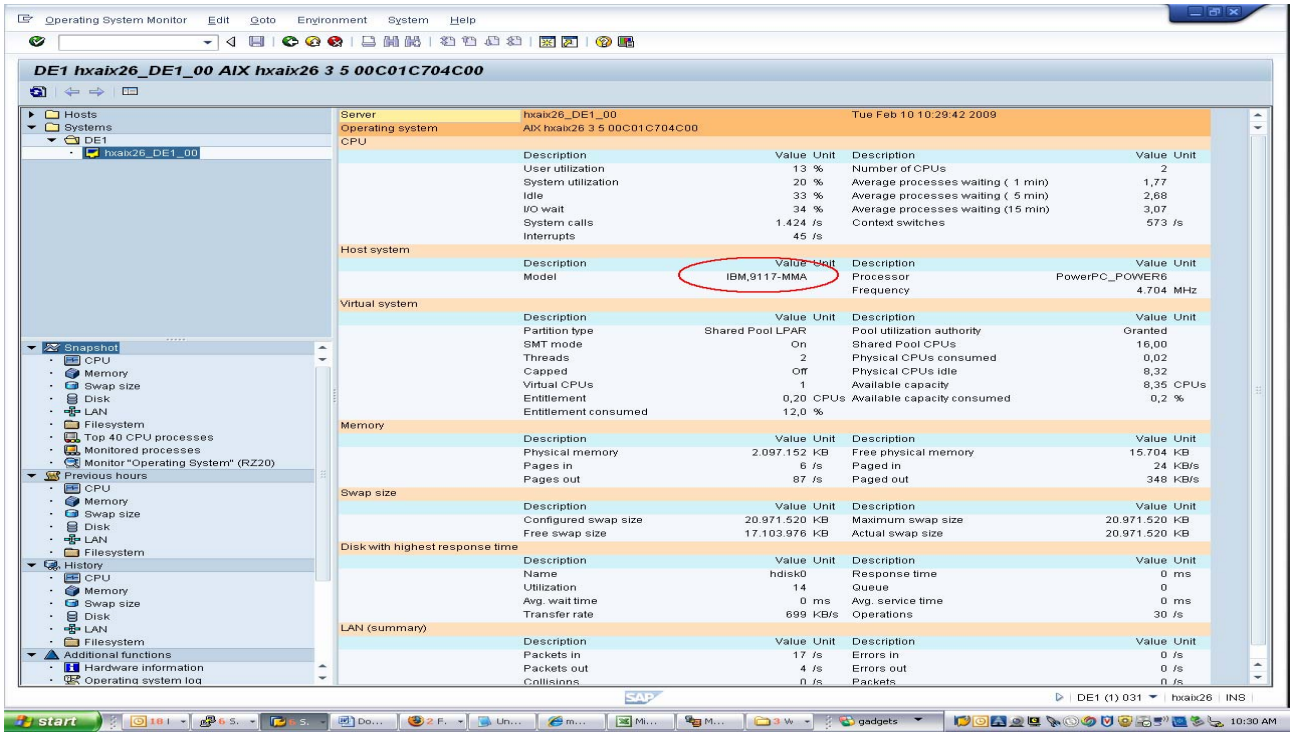
- Transaction SLICENSE in SAP system after LPM



- Requesting new licence key on the SAP Service Marketplace
- Please Note: Please thoroughly check all non-SAP third party products for hardware key problems following LPM.

SAP Verification after LPM

- Before and after the LPM is completed, the SAP Basis person should verify the hardware key changes by running transaction ST06n.



- Once the LPM is completed, the following transactions must be checked to verify a successful migration: SM21 – Check the SAP system log for any obvious errors, ST06 → OS Collector Log – check the SAP operating system collector, ST06 → Operating System Log – check the Unix log, ST04n – check the Oracle alert log for all DBMS errors

NIM and LPM

- There is a potential problem if you initiate a NIM operation on a client after a migration.
- The CPU ID of a NIM client is stored in the NIM database so that the master can perform a check that NIM client requests are coming from the same machine that was originally registered as a client.
- This id changes with a migration and subsequent NIM client requests would fail.
- As a workaround the client CPU ID check can be disabled using the fastpath "*smitty nim_cpuid_validate*".
- More details about this attribute can be found here:

  http://publib.boulder.ibm.com/infocenter/systems/index.jsp?topic=/com.ibm.aix.install/doc/insgdrf/addl_disable_client.htm&tocNode=int_8662
- This check is only relevant for client initiated NIM operations. NIM operations are typically initiated by the NIM master, in that case the physical processor validation never happens.
- Reset a NIM client with the */usr/local/bin/resetnimclient* script, before performing a NIM operation on the client. This is an Australia Post local script.

VIOS and LPAR Documentation

- If the mobile partition is to remain permanently on the destination system, the VIOS and LPAR documentation must be updated to reflect the change to the landscape.
- For example if you move an LPAR from 570-1 to 595-2 and you plan on leaving the LPAR on 595-2 indefinitely, then you must update your VIOS configuration documentation and LPAR layout diagrams.
- The UNIX team AIX/VIOS related documentation can be found here:
  - http://esbinfo:8090/display/MIO/AIX+Configuration+management
  - Reconfigure backups – Legato, mksysb, etc.

To do list.

- Test LPM with HACMP. Supported but I have not tried it yet!
- Test LPM with AMS and dual VIOS. Same as above.
- Has anyone tried these yet?

LPM Live Demo with JS22 Blades

- As long as my VPN connections works! ;-)
- JS22 Live Partition Mobility

  http://www.ibm.com/developerworks/aix/library/au-js22lpm
- Oracle DB 10gR2 Certified with IBM PowerVM Live Partition Mobility on AIX

  http://blogs.oracle.com/OIIS/2009/07/oracle_db_10gr2_certified_with.html
- Using Live Partition Mobility with SAP Instances

  https://www.sdn.sap.com/irj/scn/weblogs?blog=/pub/wlg/13150
- SAP Note 1102760 - POWER6 Live Partition Mobility:
  - The required minimum software levels are:
    - Oracle DB 10gR2 (10.2.0.4) - Single Instance only, no RAC
    - AIX 5.3 TL8 SP4, or
    - AIX 6.1 TL2 SP3
  - Please note that Live Partition Mobility is certified for single Oracle database instances only.
  - Oracle RAC is not yet supported. Formal testing with Oracle RAC is still ongoing and support will be announced at a later date.