

Live Partition Mobility

Jaqui Lynch
lynchj@forsythe.com

Presentation at:
<http://www.circle4.com/forsythe/lpm2014.pdf>



1



Why Live Partition Mobility?

Uses

- Server Consolidation
- Workload Balancing
- Can use to allow power off of unused servers to assist in Green IT efforts
- Planned CEC outages for maintenance/upgrades Impending CEC outages (e.g. hardware warning received)
- Ability to move from Power7 servers to Power8 servers (when available) without an outage. With Linux and AIX 5.3 & 6.1, you can also move from Power6 to Power7

Inactive partition migration moves a powered-off partition

- Not a crashed kernel

Partitions cannot be migrated from failed machines

- Network applications may see a brief (<2 sec) suspension toward the end of the migration, but connectivity will not be lost

IT IS NOT A REPLACEMENT FOR HACMP OR OTHER HA or DR solutions

2



POWERVM LIVE PARTITION MOBILITY (LPM)

- LPM provides the ability to move an AIX or Linux partition from one LPM-capable physical server to another *compatible* server
- What is meant by *compatible*?
 - Power Systems server requirements
 - Management console requirements
 - Virtual I/O Server (VIOS) requirements
 - Mobile partition requirements
- How are LPM servers managed?
 - HMC
 - Originally rack servers only (no Power blades)
 - Now also Flex servers
 - IVM
 - When HMC not there
 - Most Power Blades
 - SDMC
 - Supported HMC and Power Blades
 - FSM
 - Flex servers only

3



DEFINITIONS

Active Partition Mobility

Active Partition Migration is the actual movement of a running LPAR from one physical machine to another without disrupting the operation of the OS and applications running in that LPAR.

Inactive Partition Mobility

Inactive Partition Migration transfers a partition that is logically 'powered off' (not running) from one system to another.

Suspended Partition Mobility

Suspended Partition Migration transfers a partition that is suspended from one system to another.

Partition Mobility (Live or Inactive) and Partition Migration (Active or Inactive)

Refer to the same feature.

4



SUMMARY OF MIGRATION PHASES

Active

- Validate configuration
- Create new LPAR
- Create new virtual resources
- *Migrate state of LPAR in memory*
- Remove old LPAR configuration
- Free up old resources

Inactive

- Validate configuration
- Create new LPAR
- Create new virtual resources
- Remove old LPAR configuration
- Free up old resources

5



STAY CURRENT

VIOS Lifecycle

| Version | GA | EOM | EOS/EOL |
|---------|-------|-------|---------|
| 1.5 | 11/07 | 2008 | 09/11 |
| 2.1 | 11/08 | 2010 | 09/12 |
| 2.2.0 0 | 9/10 | 2011 | 09/13 |
| 2.2.1 | 10/11 | 10/12 | 04/15 |
| 2.2.2 | 10/12 | 10/13 | 09/16 |
| 2.2.3 | 4Q13 | | |
| 2.2.4 | 2Q15 | | |

Latest release:

2.2.3.3 FP27 SP03

Basically AIX 6100-09-03 -

Requires NIM at 6100-09-03 or 7100-03-03

Can be applied to 2.2.3.0, 2.2.3.1 or 2.2.3.2

Download updates from Fix Central:

<http://www-933.ibm.com/support/fixcentral/>

Download base from entitled software:

<https://www-05.ibm.com/servers/eserver/ess/ProtectedServlet.wss>

6



REQUIREMENTS FOR LPM

PLANNING IS CRITICAL

<http://www14.software.ibm.com/webapp/set2/sas/f/pm/component.html>

Hardware POWER6 Only and above

HMC v7.3.2 with MH01062

Firmware E*340_039 min

- <http://pic.dhe.ibm.com/infocenter/powersys/v3r1m5/index.jsp?topic=/p7hc3/p7hc3firmwaresupportmatrix.htm>

AIX v5.3 5300-07-01

AIX v6.1 6100-02-08

VIO Server 1.5.2.1-FP-11.1 or v2.1

RHEL5 Update 1 and SLES10 SP 1 supported (or later)

Linux LPARs must install RPMs for DLPAR

HMC v7.3.4 introduces remote migration

- Partitions can migrate between systems managed by different HMCs
- Mobility between HMC and FSM requires HMC v7r7.1.0 or later

Virtualized SAN storage (rootvg and all other vgs)

Virtualized Ethernet (SEA)

LPAR being moved cannot be using the HEA/IVE (VIO can though)

Check the prereq site:

- https://www-912.ibm.com/e_dir/eserverprereq.nsf

No dedicated anything at the time of the move

7



OTHER LPM PREREQS

Two servers – POWER6, 6+, 7, 7+, 8 (or mix thereof)

PowerVM Enterprise Edition on all servers

All LPARs on the same Open network with RMC established to HMC

LPARs must be under control of VIOS

Both systems HMC connected on the RMC network

Check VIOS levels

- Many new features require v2.1 or higher

Storage must be virtualized

- Storage must be zoned to both source and target
- No LVM based disks
- hdisks must be external and have reserve_policy=no_reserve
- See section 3.7 of the LPM red book SG24-7460

Must use Shared Ethernet Adapter

- See section 3.8 of the LPM red book SG24-7460

All resources must be shared or virtualized prior to migration

No partition at receiving server can have the same name as the LPAR being migrated from

Must have resources available at the target

- Cores, memory, adapters, etc

8



LPM

Check LPAR on HMC under Capabilities

- Look for Active and Inactive Partition Mobility Capable=True

Ensure VIO server is set up as a Mover Service Partition (MSP) under the general tab on the VIO server at each end

- By default MSP is set to no on a VIO server

Mover partition must have a VASI (Virtual Asynchronous Services Interface) device defined and configured (done automatically by HMC)

The pHypervisor will automatically manage migration of CPU and memory

Dedicated IO adapters must be de-allocated before migration

cd0 in VIO may not be attached to mobile LPAR as virtual optical device

Time of Day clocks for VIO servers should be synchronized

The operating system and applications must be migration-aware or migration-enabled

Oracle 10G supports LPM

LMB (memory region) size must be the same on both servers –
check on HMC

- Requires a whole server reboot to change

Check capability for Active Memory Sharing, suspend/resume and Trusted Boot if mobile partition is configured for them

9



CHECKLIST FROM LINDA FLANDERS AT IBM

Check processor architecture (>=POWER6)

Verify servers authorized for POWERVM

Check firmware on both servers are compatible for LPM

<http://www-01.ibm.com/support/knowledgecenter/POWER7/p7hc3/p7hc3firmwaresupportmatrix.htm>

Check LMB (logical memory block) size is the same on both servers

Verify target system has enough memory and CPU to support the mobile LPAR

Verify min_proc_units_per_virtual_proc is the same on both servers (0.05 versus 0.10) if source LPAR has 0.05 processor units

Check target server for support for AMS, suspend/resume and trusted boot if source LPAR is using any of them

10



CHECK EACH SERVER FOR PARTITION MOBILITY

https://bpichmc/hmc/wcl/T28dd

p750-Server-8233-E8B-SN069348P

General Processors Memory I/O Migration Power-On Parameters **Capabilities** Advanced

| Capability | Value |
|--|-------|
| Active Memory Sharing Capable | True |
| IBM i Capable | True |
| 5250 Application Capable | False |
| CoD Capable | True |
| Processor Capable | True |
| Memory Capable | False |
| Micro-partitioning Capable | True |
| Virtual I/O Server Capable | True |
| Logical Host Channel Adapter Capability | True |
| Logical Host Ethernet Adapter Capability | True |
| Huge Page Capable | True |
| Barrier Synchronization Register (BSR) Capable | True |
| Service Processor Failover Capable | True |
| Shared Ethernet Adapter Failover Capable | True |
| Redundant Error Path Reporting Capable | True |
| GX Plus Capable | True |
| Hardware Discovery Capable | True |
| Active Partition Mobility Capable | True |
| Inactive Partition Mobility Capable | True |
| IBM i Partition Mobility Capable | True |

OK Cancel Help

11

FORSYTHE

CHECKING IF SERVER IS MIGRATION CAPABLE

https://bpichmc/hmc/wcl/Tb1e

p740-Server-8205-E6B-SN10934CP

General Processors Memory I/O **Migration** Power-On Parameters Capabilities Advanced

Migration Capabilities

| Type | Capable | Number of Supported Migrations | Number of Migrations in progress |
|----------|---------|--------------------------------|----------------------------------|
| Inactive | True | 4 | 0 |
| Active | True | 8 | 0 |

Migration Policies

Inactive profile migration policy Partition Configuration

OK Cancel Help

https://bpichmc/hmc/wcl/Tbcd

p750-Server-8233-E8B-SN069348P

General Processors Memory I/O **Migration** Power-On Parameters Capabilities Advanced

Migration Capabilities

| Type | Capable | Number of Supported Migrations | Number of Migrations in progress |
|----------|---------|--------------------------------|----------------------------------|
| Inactive | True | 4 | 0 |
| Active | True | 8 | 0 |

Migration Policies

Inactive profile migration policy Partition Configuration

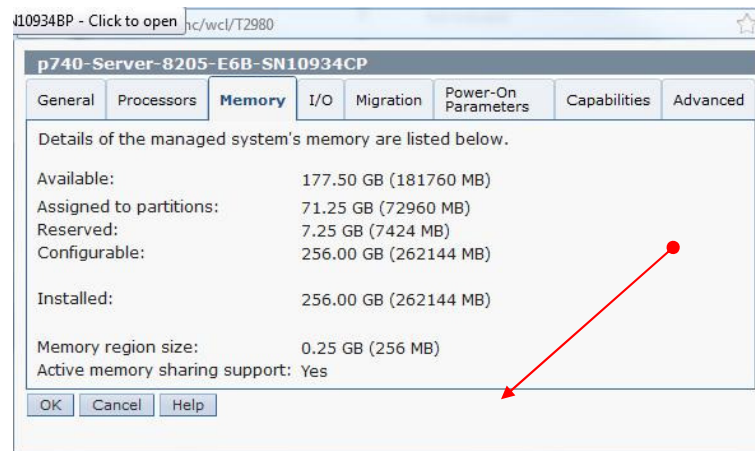
OK Cancel Help

12

FORSYTHE

Can also use
Issyscfg
Ismigrpar

CHECK LMB/MEMORY REGION SIZE ON 740

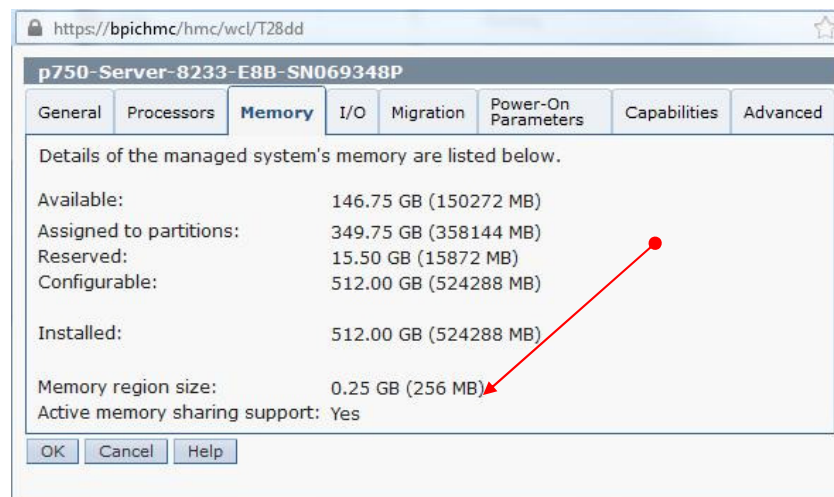


LMB size is changed using the ASMI and requires a server POWER on and off

13



CHECK LMB/MEMORY REGION SIZE ON 750



14



CHECK 740 FOR VIO MSP

https://bpichmc/hmc/content?taskId=107&refresh=185

Partition Properties - b740vio1

General Hardware Virtual Adapters Settings Other

Name: *b740vio1

ID: 1

Environment: Virtual I/O Server

State: Running

Attention LED: Off

Resource configuration: Configured

OS version: VIOS 2.2.2.3

Current profile: default

System: 8205-E6B*10934CP

☒ Mover service partition

☒ Allow performance information collection

OK Cancel Help

15



CHECK 750 FOR VIO MSP

https://bpichmc/hmc/content?taskId=105&refresh=181

Partition Properties - b750vio1

General Hardware Virtual Adapters Settings Other

Name: *b750vio1

ID: 1

Environment: Virtual I/O Server

State: Running

Attention LED: Off

Resource configuration: Configured

OS version: VIOS 2.2.2.3

Current profile: default

System: 8233-E8B*069348P

☒ Mover service partition

☒ Allow performance information collection

OK Cancel Help

16



HMC CHECKLIST FROM LINDA FLANDERS AT IBM

Server can be managed by one or more Power Systems Management Consoles (HMC, IVM, or FSM)

- Redundant configuration supported
- Mobility between HMCs and FSM requires HMC V7R7.1.0 or later
- No mobility between IVM and either HMC or FSM

Minimum HMC V7R3.2 or minimum HMC V73.4 for *remote* HMC operation

- Remote = each server managed by a different console
- Configure SSH key authentication between the remote consoles

Functional RMC daemons

Network access to involved partitions (VIOS partitions, mobile LPAR), server, and remote management console

17



REQUIREMENTS FOR REMOTE MIGRATION

Ability to use LPM between 2 servers on different HMCs

A local HMC managing the source server

A remote HMC managing the target server

Functional RMC daemons

Version 7.3.4 or later of the HMC software

Network access to the remote HMC

SSH key authentication to the remote HMC and all involved LPARs (VIOS and actual LPAR)

Plus all the other requirements for single HMC migration

If FSM to HMC or vice versa then need HMC v7.7.1.0 or higher

18



VIOS REQUIREMENTS CHECKLIST

VIOS on both source and destination server must provide:

- Access to the same network (VLAN and subnet) with a Shared Ethernet adapter
 - Validation also checks virtual switch name and VEB/VEPA configuration
- Access to the same physical storage either from:
 - External Fibre Channel storage system using virtual SCSI or NPIV
 - External iSCSI storage using virtual SCSI

Virtual storage:

- VIOS virtual adapters cannot be marked as required and should not be marked for “Any client”
- Destination VIOS must have enough “available” virtual adapters

Other:

- At least one VIOS per server must be configured as an MSP
- For shared memory LPARs, destination must have a paging device available
- Consider configuring a VIOS as a time reference partition

MOBILE PARTITION REQUIREMENTS CHECKLIST 1/2

Check mobile partition configuration for:

Operating system support
 Functioning RMC daemons
 No huge memory pages*
 No BSR arrays*
 No workload group
 No redundant error reporting*
 No open consoles (warning only)
 Storage on external storage unit and accessible by both VIOS LPARs
 Check that LPAR name will be unique on destination server
 Check MAC address will be unique on destination server
 Not a service partition
 No physical I/O*
 No logical ports configuration on Integrated Virtual Ethernet (IVE)/Host Ethernet Adapter (HEA)
 AIX 6100-05 (or higher) clients can convert LHEA to virtual Ethernet during migration
 Verify valid processor compatibility mode
 Virtual network configuration compatible with destination VIOS

* Okay for inactive migrations

MOBILE PARTITION REQUIREMENTS CHECKLIST 2/2

Also check configuration for:

If mobile LPAR using AME (Active Memory Expansion)
Make sure it's supported on the destination server

If the mobile partition is suspend-resume capable, make sure the target has a reserved storage pool greater than or equal to 110 percent of the lpar size

Migrating IBM i LPAR

1. Verify the destination server supports the migration of IBM i mobile partitions and the restricted I/O mode
2. Verify the IBM i mobile partition is in the restricted I/O mode

If shared processor pool with an entitlement below 0.1 processing units and greater than or equal to 0.05 (POWER7+), make sure the destination server also shares that capability

LUNs using NPIV need to be mapped to both world wide port names on each client Virtual Fibre Channel adapter

All I/O resources must be shared or virtualized prior to migration – dedicated devices need to be removed

21

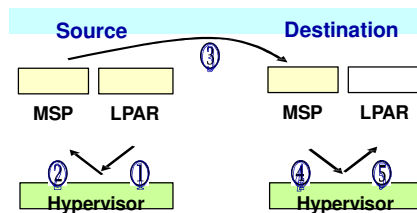


MSP USED FOR ACTIVE MIGRATIONS

- In active migrations, in addition to moving the LPAR configuration on the HMC, the partition memory state is moved

• Memory state includes:

- Partition's memory
- Hardware page table (HPT)
- Processor state
- Virtual adapter state
- Non-volatile RAM (NVRAM)
- Time of day (ToD)
- Partition configuration



22



VIRTUAL ASYNCHRONOUS SERVICES INTERFACE (VASI)

- The VASI device provides a communication path between the Virtual I/O Server and the Hypervisor
 - Enabled when VIOS MSP is enabled
- Source and destination MSPs use a VASI device to gain access to a mobile partition's state during an active migration
- Additional VASI devices are added to a VIOS when Active Memory Sharing is configured
- View VASI devices in the VIOS CLI:


```
$ lsdev -virtual | grep vasi
vasi0    Available  Virtual Asynchronous Services
          Interface (VASI)
```
- Activity of VASI device can be monitored with VIOS **vasistat** command

23



RESOURCE MONITORING AND CONTROL (RMC)

- RMC provides a generalized, distributed framework for managing resources within a single system, multiple systems over a network, or a cluster
- RMC purpose:
 - Allows management console to communicate with partitions to perform operations such as shutdown, DLPAR, service event reporting, and virtual device management
 - Used by cluster management tools to monitor, query, modify, and control cluster resources
- LPM utilizes RMC for migrations
 - RMC operates over the network between the management console(s), the VIOS partitions, the MSP partitions, and the mobile LPAR partition

24



CHECK RMC CONNECTIONS

- Verify RMC is active for: MSPs, VIOS partitions, mobile LPAR, and HMC/FSM
- To quickly check the status between LPAR and management console, run the **Operations -> Shut Down HMC/FSM** task
 - If OS options are grayed out (unavailable), the RMC connection is not active between the LPAR and the management console

RMC connections take 5-7 minutes to establish, so wait after a reboot before testing

Use `rmcctrl` command to stop and restart RMC subsystem if necessary

Can take up to 7 minutes for daemons to resynchronize

Use `recfgct` to reinitialize RMC if necessary

Last resort reboot the HMC/FSM

RMC is used by LPM for migrations and operates over the network between management consoles, VIOS partitions, MSP partitions and the mobile LPAR

If all else fails reboot the HMC or FSM

25



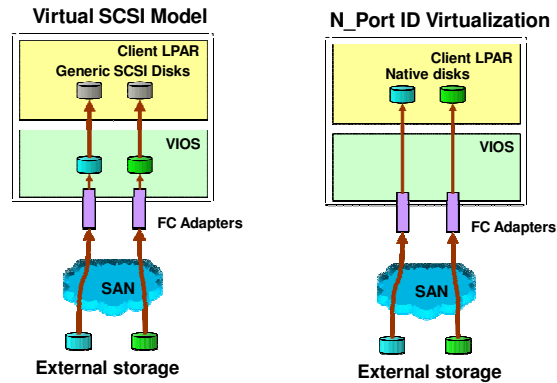
CONFIGURE EXTERNAL STORAGE

- Storage devices can be provided to the mobile partition via:
 - Virtual SCSI adapters or virtual Fibre Channel adapters with NPIV
 - Includes storage from VIOS Shared Storage Pool (SSP) as long as source and destination are part of same VIOS storage group
- Virtual SCSI
 - Storage must be external physical volumes accessible to involved VIOS partitions via Fibre Channel SAN or iSCSI SAN
 - Zone same LUNs to both VIOS LPARs HBA ports
 - Disable SCSI reservations on the hdisks PRIOR to creating virtual target
- Virtual Fibre Channel
 - Storage must be external and accessible via FC ports on the involved VIOS partitions
 - Have to zone virtual WWPNS

26



THE CONTEXT: POWERVM DISK ACCESS



This is relevant for LPM as for:

- Zoning
- LUN-masking

27



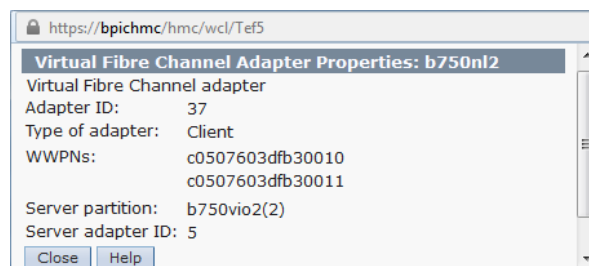
ZONING AND MASKING FOR NPIV

Client Virtual Fibre channel adapters have two WWPNs
Need to zone both WWPNs below for LPM

The migration creates adapters on the target VIOS and removes adapters on the source VIOS after the migration

The migration maps a physical FC port on the VIOS to the vfchost adapter

Management console attempts to use same adapter IDs if available



28



VIRTUAL NETWORKING

- Mobile partition must use a virtual Ethernet adapter for network access
 - Destination VIOS must have an SEA that supports LPAR's subnet and VLAN
- SR-IOV logical port is a physical device and not compatible with LPM
- As of AIX 6.1 TL5, mobile LPARs can use HEA port, however:
 - HEA port must be part of an Etherchannel configuration with virtual Ethernet adapter configured as backup
 - Etherchannel fails over to virtual Ethernet adapter, then Etherchannel and HEA ports are removed
 - Uses "**smit migration**" fastpath on mobile LPAR

29



PROCESSOR COMPATIBILITY MODE

- A partition's *preferred* mode is configured in partition profile
- A partition's *current* mode is what the Hypervisor is using
 - May be different from the preferred mode if it was not available
- Destination server must support the mobile LPAR's *current* processor compatibility mode for active LPM

Processor compatibility modes supported by server type:

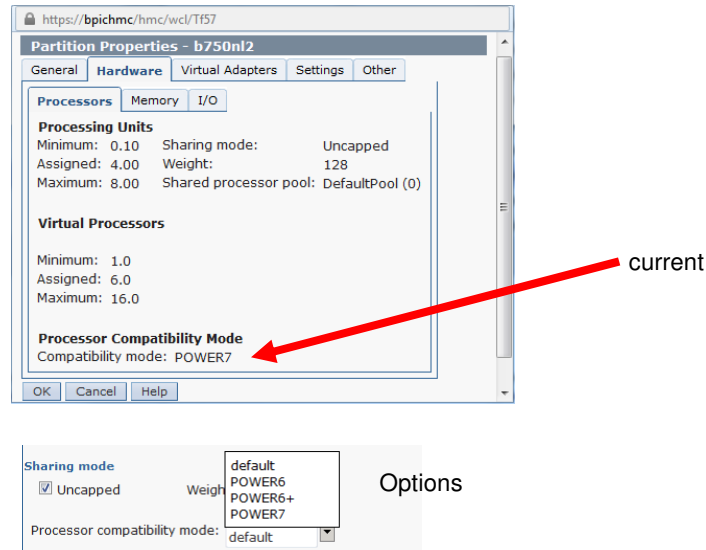
| Server Processor Type | Supported Current Modes | Supported Preferred Modes |
|-----------------------------------|---|--|
| POWER6 | POWER5, POWER6, POWER6 enhanced | default, POWER6, POWER6 enhanced |
| Refreshed POWER6 server (POWER6+) | POWER5, POWER6, POWER6+, POWER6+ enhanced | default, POWER6, POWER6+, POWER6+ enhanced |
| POWER7 and POWER7+ | POWER5, POWER6, POWER6+, POWER7 | default, POWER6, POWER6+, POWER7 |
| POWER8 | POWER6, POWER6+, POWER7, POWER8 | Default, POWER6, POWER6+, POWER7, POWER8 |

There is no POWER7+ mode. POWER7 and POWER7+ use same instruction set.

30



PROCESSOR COMPATIBILITY MODE

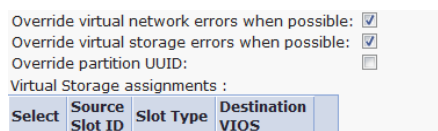


31



DUAL (REDUNDANT) VIOS PARTITIONS

- The mobile partition will attempt to have the same configuration on the destination server
- If the mobile partition is using client mirroring across dual VIOS partitions, the mobile partition can migrate to dual or single VIOS configuration
- If the mobile partition is using MPIO across dual VIOS partitions, then it must migrate to dual VIOS configuration
 - Workaround: remove one path, then migrate partition, or use override option during migration (shown below)
 - Result is the mobile partition will use only one VIOS on destination



32



INACTIVE MIGRATION

- Differences in configuration for inactive migrations:
 - LPAR may be configured with huge memory pages and barrier synchronization registers
 - Any physical I/O configured in the profiles will be removed during the migration
 - The preferred processor compatibility mode must be supported by the destination server
 - Mobile partition OS versions that support virtual devices and POWER6 servers but are not supported for active migrations may be supported for inactive migrations
- Either the last activated partition profile or the last running configuration is used, therefore a partition that has *never* been activated cannot be migrated
 - Workaround for new LPARs: Activate to SMS then shut down

33



COMMAND LINE INTERFACE -- MIGRLPAR

- Migrations can be scripted using the migrldpar command
- Run migrldpar --help or man migrldpar to see the syntax (it's updated as features are added)
- You can be very specific in selecting the source and destination MSP servers, the mapping of virtual SCSI and virtual FC servers and slots, and other configuration details

34



VALIDATION – HIGH LEVEL

- Capability and compatibility check
- Resource Monitoring and Control (RMC) check
- Partition readiness
- System resource availability
- Virtual adapter mapping
- Operating system and application readiness check

35



LPM – VALIDATION PHASE

Checks the source and destination systems, POWER Hypervisor, Virtual I/O, Servers, and mover service partitions for active partition migration capability and compatibility

Checks that the RMC connections to the mobile partition, the source and destination Virtual I/O Servers, and the connection between the source and destination mover service partitions are established

Checks that there are no required physical adapters in the mobile partition and that there are no required virtual serial slots higher than slot 2

Checks that no client virtual SCSI disks on the mobile partition are backed by logical volumes and that no disks map to internal disks

Checks the mobile partition, its OS, and its applications for active migration capability. Checks that the logical memory block size is the same on the source and destination systems Checks that the mobile partition is not configured with barrier synchronization registers Checks that the mobile partition is not configured with huge pages

Checks that the partition state is active or running

Checks that the mobile partition is not in a partition workload group

Checks the uniqueness of the mobile partition's virtual MAC addresses

Checks that the mobile partition's name is not already in use on the destination server

Checks the number of current active migrations vs the number of supported active migrations

36



B750NL2 LPAR INITIAL VALIDATION SCREEN

https://bpichmc/hmc/content?taskId=99&refresh=163

Partition Migration Validation - p750-Server-8233-E8B-SN069348P - b750nl2

Fill in the following information to set up a migration of the partition to a different managed system. Click Validate to ensure that all requirements are met for this migration. You cannot migrate until the migration set up has been verified.

Source system : p750-Server-8233-E8B-SN069348P
 Migrating partition: b750nl2
 Remote HMC:
 Remote User:
 Destination system: p740-Server-8205-E6B-SN10934CP
 Destination profile name:
 Destination shared processor pool:
 Source mover service partition:
 Destination mover service partition:
 Wait time (in min):
 Override virtual network errors when possible: ☐
 Override virtual storage errors when possible: ☐
 Virtual Storage assignments :

| Select | Source Slot ID | Slot Type | Destination VIOS |
|--------------------------|----------------|-----------|------------------|
| <input type="checkbox"/> | | | |

37



B750NL2 AFTER CLICK ON VALIDATE

https://bpichmc/hmc/wcl/Tdc5

Partition Migration Validation - p750-Server-8233-E8B-SN069348P - b750nl2

Fill in the following information to set up a migration of the partition to a different managed system. Click Validate to ensure that all requirements are met for this migration. You cannot migrate until the migration set up has been verified.

Source system : p750-Server-8233-E8B-SN069348P
 Migrating partition: b750nl2
 Remote HMC:
 Remote User:
 Destination system: p740-Server-8205-E6B-SN10934CP
 Destination profile name: default
 Destination shared processor pool: DefaultPool (0)
 Source mover service partition: b750vio2
 Destination mover service partition: b740vio2
 Wait time (in min):
 Override virtual network errors when possible: ☒
 Override virtual storage errors when possible: ☒
 Override partition UUID: ☐
 Virtual Storage assignments :

| Select | Source Slot ID | Slot Type | Destination VIOS |
|-------------------------------------|----------------|-----------|------------------|
| <input checked="" type="checkbox"/> | 27 | SCSI | b740vio2 |
| <input checked="" type="checkbox"/> | 57 | SCSI | b740vio2 |
| <input checked="" type="checkbox"/> | 37 | Fibre | b740vio2 |
| <input checked="" type="checkbox"/> | 47 | Fibre | b740vio2 |

38



MIGRATION

A screenshot of a web browser window displaying the BMC Helix console. The address bar shows the URL: https://bpmchmc/hmc/content?taskId=50&refresh=106. The main content area has a dark blue header with the text "Partition Migration Status : b750n12". Below this, the text "Migration status :" is followed by a light blue box containing the text "Migration Complete". Below this box is a "Stop..." button. At the bottom, a progress bar is shown with the text "Progress (%): 100".

[illegible]

MIGRATION STEPS 1/3

- If validation passes, migration can begin
- From this point, all state changes are rolled back if an error occurs
- HMC creates a shell partition on the destination system
- HMC configures the source and destination Mover Service Partitions (MSP)
 - MSPs connect to PHYP thru the Virtual Asynchronous Serial Interface (VASI)
- MSPs set up a private, full-duplex channel to transfer partition state data
- HMC sends a Resource Monitoring and Control (RMC) event to the mobile partition so it can prepare for migration
- The HMC creates the virtual target devices and virtual SCSI adapters in the destination MSP
 - You do NOT do this manually; it must happen under Hypervisor control
- The MSP on the source system starts sending the partition state to the MSP on the destination server



MIGRATION STEPS (2/3)

- The source MSP keeps copying memory pages to the target in successive phases until modified pages have been reduced to near zero
- The MSP on the source instructs the PHYP to suspend the mobile partition
- The mobile partition confirms the suspension by suspending threads
- The source MSP copies the latest modified memory pages and state data
- Execution is resumed on the destination server and the partition re-establishes the operating environment
- The mobile partition recovers I/O on the destination server and retries all uncompleted I/O operations that were going on during the suspension
 - It also sends gratuitous ARP requests to all VLAN adapters

43



MIGRATION STEPS 3/3

- When the destination server receives the last modified pages, the migration is complete
- In the final steps, all resources are returned to the source and destination systems and the mobile partition is restored to its fully functional state
- The network channel between MSPs is closed
- The VASI channel between MSP and PHYP is closed
- VSCSI and VFC adapters on the source MSP are removed
- The HMC informs the MSPs that the migration is complete and all migration data can be removed from their memory tables
- The mobile partition and all its profiles are deleted from the source server
- You can now add dedicated adapters to the mobile partition via DLPAR as needed, or put it in an LPAR workload group

44



EFFECTS OF MIGRATION 1/2

- Server properties
 - The affinity characteristics of the logical memory blocks may change
 - The maximum number of potential and installed physical processors may change
 - The L1 and/or L2 cache size and association may change
 - ✓ This is not a functional issue, but may affect performance characteristics
- Console
 - Any active (serial) console sessions will be lost when the partition is migrated
 - Console sessions must be re-established on the target system by the user after migration
- Network
 - A temporary network outage of seconds is expected to occur as part of suspending the partition
 - ✓ Temporary network outages may be visible to application clients, but it is assumed that these are inherently recoverable

45



EFFECTS OF MIGRATION 2/2

- VSCSI Server and Server Virtual Fibre Channel Adapters
 - Adapters that are configured with the remote partition set to the migrating partition will be removed
 - ✓ Adapters that are configured to allow “any” partition to connect will be left configured after the migration
 - ✓ Any I/O operations that were in progress at time of the migration will be retried once the partition is resumed
 - As long as unused virtual slots exist on the target VIO server, the necessary controllers and target devices will be automatically created
- Error logs
 - All of the error messages that the partition received will appear on the target system
 - Error logs contain the machine type, model and serial number so it is possible to correlate the error with the system that detected it
- Partition time
 - When a partition is migrated the Time of Day and timebase values of the partition are migrated.

46



WHAT'S NEW WITH VIOS 2.2.2.0 (4Q 12)

Several new functions added with VIO 2.2.2.0 specific to LPM

Up to 3x faster

Previous versions of code used a single-threaded process to copy memory from source to target VIO servers (Mover Service Partitions)

New improvements in single session LPM performance can accelerate mobility for a single session by up to 3X over previous releases

To see this improvement one would need a fast(er) network between source and destination VIO servers – namely an aggregate/etherchannel or 10Gbe

New Firmware Level support matrix Number of concurrent

migrations

New pseudo device attributes like concurrency level Specify ports using

`vios_fc_port_name`

New error codes

47



NEW WITH VIOS 2.2.2.0 – CONCURRENT MIGRATIONS

Increased number of concurrent migrations

IBM PowerVM and the Hardware Management Console now support up to 16 concurrent LPM activities.

710_xxx and earlier firmware levels restricted to 4 concurrent LPM operations

720_xxx and 730_xxx are restricted to 8 concurrent LPM operations

740_xxx and higher can support up to 16 concurrent LPM operations

All concurrent LPM operations must have the same source and destination managed systems

Mover Service Partitions (MSP) with 1 Gb/s network adapters are restricted to a maximum of 4 concurrent LPM operations

To support more than 8 concurrent LPM operations, you must have more than one pair of VIO partitions

Systems managed with IVM support up to 4 concurrent migrations

48



NEW WITH VIOS 2.2.2.0 – PSEUDO DEVICE

New pseudo device to support LPM tuneables

The pseudo device **violpm0** is created by default when VIO server version 2.2.2.0 or higher is installed

Device attributes for **violpm0** can be used to control live partition mobility operations

Use normal lsdev/chdev padmin commands to query/change attributes

Example

```
$ lsdev -dev vioslpm0 -attr
$ chdev -dev vioslpm0 -attr cfg_msp_lpm_ops=5
```

49



NEW WITH VIOS 2.2.2.0 – VIOS_FC_PORT_NAME

Specify fibre channel ports using **vios_fc_port_name**

Run the **lslparmigr** command to show a list of available slot IDs for a VIOS partition

Run the **migrpar** command to accomplish the following tasks:

- Specify virtual slot IDs for one or more virtual adapter mappings

- Validate the specified slot IDs

Note: You can specify the port name of the Fibre Channel to be used for creating Fibre Channel mapping on the source server when you are performing partition migration.

You can use the HMC command line interface to specify the port name.

List all the valid port names of the Fibre Channel by running the **lsnports** command.

From the list of the valid port names, specify the port name that you want to use by running the **migrpar** command with the attribute **vios_fc_port_name** specifying the port name you want to use.

50



NEW WITH VIOS 2.2.2.0

The following attributes of pseudo device can be modified by using the **migr1par** command:

num_active_migrations_configured
concurr_migration_perf_level

Run the following HMC command to modify the attribute values of the pseudo device, for example to set the number of active migrations to 8 run:

```
migr1par -o set -r lpar -m <CecName> -p <lparName> -i
"num_active_migrations_configured=8"
```

The default value for this attribute is 4. To run the maximum number of supported partition mobility operations on the Virtual I/O Server (VIOS), this value must be set the supported maximum number.

To set the amount of resources allocated for each mobility operation to a value of 4, run the following command:

```
migr1par -o set -r lpar -m <CecName> -p <lparName> -i
"concurr_migration_perf_level=4"
```

51



TUNING LPM

- As of VIOS 2.2.2:
 - Start of performance enhancements
 - More concurrent operations possible
 - Ability to choose destination FC port from **migr1par** command
 - **vioslpm0** device has attributes that can be tuned

```
$ ioslevel
2.2.3.3
$ lsdev -dev vioslpm0 -attr
```

| attribute | value | description | user_settable |
|-----------------|-------|---|---------------|
| cfg_msp_lpm_ops | 8 | Configured number of concurrent LPM operations for this MSP | True |
| concurrency_lvl | 3 | Concurrency level | True |
| lpm_msnap_succ | 1 | Create mini-snap for successful migrations | True |
| max_lpm_vasi | 1 | Maximum number of VASI adapters used for LPM operations | False |
| max_vasi_ops | 8 | Maximum number of concurrent LPM operations per VASI | False |
| tcp_port_high | 0 | TCP highest ephemeral port | True |
| tcp_port_low | 0 | TCP lowest ephemeral port | True |

Concurrency level: values are 1-5, default is 3, higher values use less resources

LPM migration snap success: Should traces be saved for successful operations?
 1=yes; Support can analyze for performance issues

TCP ports: low and high values set the range of ports to use; 0=any

52



POWERVM 2.2.3

Announce 10/8/2013 GA 11/29/2013

- Shared Storage Pool Enhancements (SSP 4)
 - Mirrored pool – redundancy improvement
 - Pool Shrinking Improvement – Elasticity
 - Ability to remove a LUN from the SSP
 - Dynamic disk growth – Elasticity
 - New simpler SSP commands (failgrp, pv and lu commands)
 - Scale improvements
 - 16 nodes, 1024 physical disks, 1000 LPARs
- Import existing storage into SSP (lab services)
- Live Partition Mobility Performance Improvements
- vSCSI Redundancy Improvements
- Improvements in VIOS Performance Advisor
 - New advice for Fiber channel, SEA and SSP
- Enterprise Edition includes NEW PowerVP virtualization performance monitor
- Ability to evacuate a server with one HMC command

53



VIO 2.2.3 SEA CHANGES

Traditional SEA setup

Ent4 is the virtual adapter defined at the HMC with external access
(SEA goes here)

VIO1 is priority 1 and VIO2 is priority 2

Ent5 is the virtual adapter on Vlan 1 with no external
(IP will go here)

Ent6 is the control channel on vlan 255

`mkvdev -sea ent0 -vadapter ent4 -default ent4 -defaultid 1 -attr ha_mode=auto ctl_chan=ent6`
Creates ent7 as the SEA

Do not mess up priorities or ctl_chan or you will cause a spanning tree loop

Update with 2.2.3

See chapter 4 of SG248198- Redbook on 2.2.3 Enhancements

SEA setup has been simplified

Requirement removed for dedicated control channel and VLAN ID for each SEA failover configuration

Multiple SEA pairs can now share VLAN 4095 within the same virtual switch and no ctl_chan is needed

HMC (>= 7.8) reserves 4095 for internal management traffic

Requires VIOS 2.2.3, HMC 7.7.8 and firmware 780 or higher

54



SERVER EVACUATION COMMAND

- As of HMC R7V780+ and VIOS 2.2.3, you can migrate all* partitions from one server to another
 - All* as long as another maximum is not hit (max for HMC or MSP)
 - They will merge with any existing partitions already on target
- HMC command:


```
migr1par -o m -m source_server -t target_server -all
```
- Make sure target has enough resources to support all LPARs
- For performance, have a big Ethernet pipe and lots of CPU:
 - At least 3-4 processors and 3-4 VPs

55



NEW WITH HMC 7.7.8 (NOVEMBER 2013)

New --all flag added to the migr1par command

Single command to migrate all AIX, Linux and IBM i partitions between two managed systems
There can only be a single source and single destination system specified

```
migr1par -o m -m <SourceServer> -t <DestinationServer> --all
```

Easy command to stop migration of all AIX, Linux and IBM i partitions

```
migr1par -o s -m <SourceServer> --all
```

56



HMC v8

Required for POWER8

Runs on cr5 or C08 or higher

Will not run on earlier HMCs

Validates entitlement for POWER8

Introduces new Performance and Capacity Monitoring Task

Provides reports on resource utilization

NIST support – updates to JVM

LPM improvements to vSCSI performance

SR-IOV support

Dynamic partition remote restart can be changed when LPAR deactivated,
not just at creation time

Absolute values for DLPAR

DOES NOT SUPPORT ANYTHING PRIOR TO POWER6

57



POWER VC AND POWER VP

Power VC

- Provides virtualization management
- Virtual image management and deployment
- Resource pooling and dynamic VM placement
- Openstack based
- Ongoing optimization and VM resilience

Power VP (GA 11/29/2013)

- Realtime graphical performance monitor
- Included with Power VM Enterprise edition – can be purchased alone
- Provides performance information for virtualized systems
- Start with topology view then drill down into hardware and then LPARs
- Uses system wide and partition collectors
- Requires firmware 7.7 or higher
- Not available on all models – see announcement for list
- Information can be found at:
<http://www.ibm.com/systems/power/software/Performance/>

58



REFERENCES

LPM Preparation Checklist

<http://www.redbooks.ibm.com/abstracts/tips1185.html?Open>

LPM Setup Checklist

<http://www.redbooks.ibm.com/abstracts/tips1184.html?Open>

PowerVM Enhancements

<http://www.redbooks.ibm.com/redbooks/pdfs/sg248198.pdf>

SG24-7940 - PowerVM Virtualization - Introduction and Configuration

<http://www.redbooks.ibm.com/redbooks/pdfs/sg247940.pdf>

SG24-7590 – PowerVM Virtualization – Managing and Monitoring

<http://www.redbooks.ibm.com/redbooks/pdfs/sg247590.pdf>

Upcoming Forsythe Talks

<http://www.circle4.com/forsythetalks.html>

Movie replays

<http://www.circle4.com/movies>

59



THANK YOU FOR YOUR TIME



If you have questions please email me at:

lynchi@forsythe.com

Presentation at:

<http://www.circle4.com/forsythe/lpm2014.pdf>

60

