

Partitioning AIX 101

Jaqui Lynch

<http://www.circle4.com/jaqui/papers/aixpartition.pdf>

Jaqui.lynch@mainline.com



Agenda

- Partitioning Concepts
- Hardware
- Software
- Planning
- Hints and Tips
- References



Partitioning Concepts

- Logical Partitions
- Full System Partition
- Affinity Partitions
 - CPU and memory resources are allocated in fixed patterns based on multi-chip module (MCM) boundaries
- Managed Systems
 - Partitions
- Profiles
 - Partition Profiles
 - System Profiles



Dynamic LPAR

- Add processors to partition
- Move processors between partitions
- Remove processors from a partition
- Add memory to a partition
- Move memory from one partition to another
- Remove memory from a partition
- Add a PCI adapter
- Move a PCI adapter
- Remove a PCI adapter



Reasons to Partition

- Consolidation
- Production and Test on same hardware
- Multiple Operating Systems
- Consolidate Applications on different time zones
- Complying with license agreements



Role of the HMC

- Required to partition any box
- Can use HMC to manage systems
- Provides a console to manage hardware
- Detecting, reporting and storing changes in hardware
- Service focal point (requires Ethernet)
- Vterms to partitions



HMC in an SP Cluster 1600

- Shows CWS the LPAR definitions
- Provides hardware status to CWS
- Manages LPAR definitions
- Connects via rs232 to p690
- Connects via Ethernet to CWS



Notes on the HMC

- Required at all times by p655, p670 and p690 regardless of whether they are partitioned
- On some p670 and p690 it may have been ordered as a feature code (#7316)
- 8 port and 128 port async adapters
 - Only 2 serial ports so get at least one of these
- Can order extra Ethernet card for private network if desired



HMC Rules

- Taken from p690 Technical Support Certification Guide
- The HMC provides two integrated serial ports.
 - One serial port required per pSeries server
 - One serial port required for modem attachment if the Service Agent Call-Home function is implemented
- 8- or 128-port asynchronous adapter (FC 2943 or 2944) should be used to extend (maximum of 2).
- The first HMC (FC 7315) that was announced with the p690 supported up to four managed systems.
- pSeries 690 can be attached to two separate HMCs for redundancy.
- The 128-port adapter, in combination with a Remote Asynchronous Node, can be used for a long-distance solution between HMC and a managed system. This will provide distances of up to 1100 feet or 330 meters, while normal RS-232 connections allow up to 15 meters.
- An Ethernet connection between the HMC and each active partition on the partition-capable pSeries server is required.
 - This connection is utilized to provide several system management tasks, such as dynamic logical partitioning to each individual partition, and collection and passing of hardware service events to the HMC from the partition for automatic notification of error conditions to IBM.



Hardware

| Product | Max Procs | Max GB Memory | Max I/O Drawers | Max Partitions |
|----------|----------------|---------------|-----------------|----------------|
| P690 | 32 (16 on HPC) | 256 | 8 | 16 |
| P670 | 16 | 128 | 3 | 16 |
| P655 651 | 8 | 32 | 1 | 2 |
| P650 6m2 | 8 | 64 | 8 | 8 |
| P630 6c4 | 4 | 32 | 2 | 4 |
| P630 6e4 | 4 | 32 | 0 | 2 |



Supported Operating Systems

- AIX 5.2
- Aix 5.1
 - Does not support:
 - Dynamic LPAR
 - Memory Capacity Upgrade on Demand
 - Dynamic Processor Sparing
 - Dynamic CPU Guard
- Suse Linux
- No version of AIX prior to v5 will work
- Check required ML levels for each box



Software

- Make sure HMC and all boxes are at the 10/2002 microcode level
- pSeries Microcode can be found at:
 - <https://techsupport.services.ibm.com/server/nav?fetch=hm>
- HMC Corrective Service can be found at:
 - <https://techsupport.services.ibm.com/server/hmc//corrsv.html>
- Latest HMC Software version is R3v2.2



Planning

- Each LPAR must have the following
 - 1 processor
 - 256mb memory
 - 1 boot disk
 - 1 adapter to access the disk
 - 1 Ethernet adapter to access the HMC
 - An installation method such as NIM
 - A means of running diagnostics



Memory

- In full system partition mode all memory is allocated to the system
- In LPAR mode some memory is reserved for LPAR use
 - Hypervisor - 256mb
 - TCE (Translation Control Entry) – 256mb
 - Used to translate I/O addresses to system memory addresses
 - Page Table Entries (min 256mb)
 - So overhead for the first 256mb partition is 768mb
- For 2 or more LPARS expect overhead to be at least 2gb memory

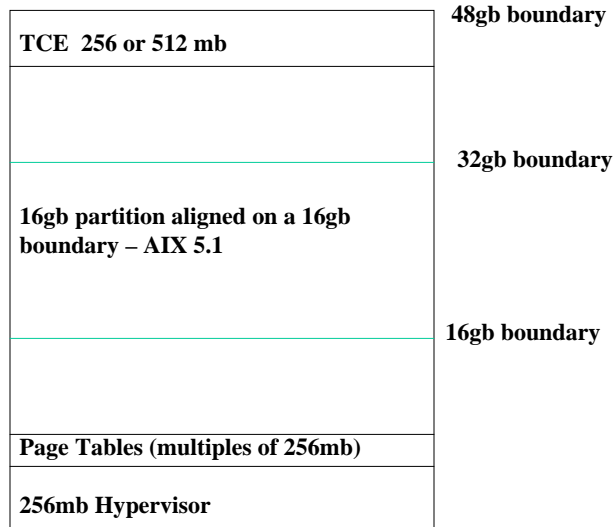


Real Mode Address Region (RMO)

- Small Real Mode Address Region
 - Allows you to use managed system memory more efficiently
 - Only valid for Linux and AIX 5.2
 - Avoids the memory boundary constraints
 - AIX 5.1 may not boot if you turn this on
- Large Real Mode Address Region
 - Assigns memory on 256mb, 1gb and 16gb boundaries (contiguous real mode memory)
 - Partition $\leq 16\text{gb}$ gets 1gb plus the rest in 256mb increments



Memory



More on Memory

- Hypervisor fixed at 256mb at address 0
- TCE (top of memory)
 - I/O and DMA translation
 - 256mb for first I/O drawers
 - 512mb if five or more drawers
 - Allocated at partition activation
- Page Tables
 - Allocated at partition activation
 - One per partition
 - 1/64th memory for the partition rounded to n*2
 - 1.5gb partition needs 24mb but rounds to 32mb



Memory 1/2

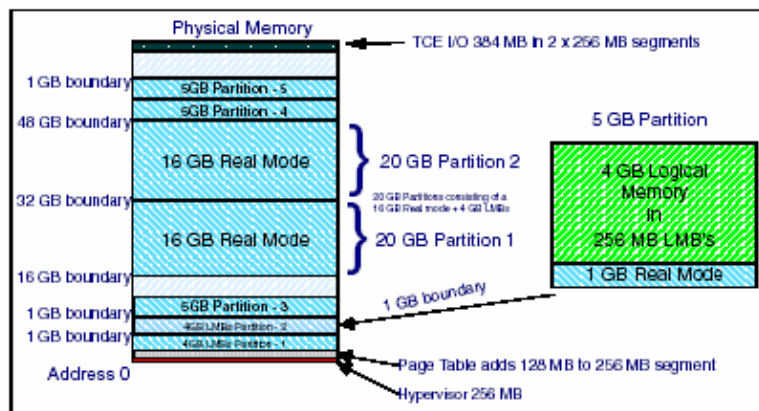


Figure 4-1 Successful allocation of five partitions

Courtesy IBM eServer Certification Study Guide – p690 Technical Support Pg 96



Memory 2/2

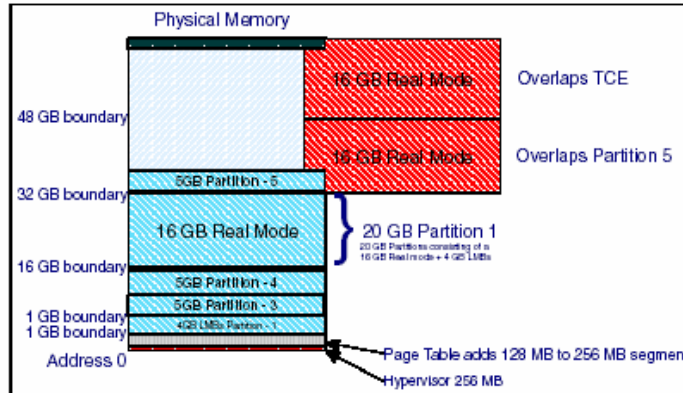


Figure 4-2 Unsuccessful allocation of five partitions

Courtesy IBM eServer Certification Study Guide – p690 Technical Support Pg 96



P690 Reserved Memory

| Total Mem GB | Approx Overhead GB | Approx Usable | Max Part pre 10/02 AIX/Lix | Max Part Post 10/02 AIX 5.1 | Max Part Post 10/02 AIX 5.2/Lix |
|--------------|--------------------|----------------|----------------------------|-----------------------------|---------------------------------|
| 4 | .75 to 1 | 3 to 3.25 | 3 and 0 | 13 and 0 | 13 |
| 8 | .75 to 1 | 7 to 7.25 | 6 and 0 | 16 and 0 | 16 |
| 16 | .75 to 1 | 15 to 15.25 | 14 and 0 | 16 and 0 | 16 |
| 24 | 1 to 1.25 | 22.75 to 23 | 16 and 0 | 16 and 0 | 16 |
| 32 | 1 to 1.25 | 30.75 to 31 | 16 and 0 | 16 and 0 | 16 |
| 48 | 1.25 to 1.75 | 46.25 to 46.75 | 16 and 1 | 16 and 1 | 16 |
| 64 | 1.5 to 2 | 62 to 62.5 | 16 and 2 | 16 and 2 | 16 |
| 96 | 2 to 2.5 | 93.5 to 94 | 16 and 4 | 16 and 4 | 16 |
| 128 | 2.5 to 3.5 | 124.5 to 125.5 | 16 and 6 | 16 and 4 | 16 |
| 192 | 3.5 to 4.5 | 187.5 to 188.5 | 16 and 10 | 16 and 10 | 16 |
| 256 | 5 to 6 | 250 to 251 | 16 and 14 | 16 and 14 | 16 |



Hints and Tips

- Large Real Mode Regions
 - Start all partitions >16gb before smaller ones
 - If all partitions >16gb then start the largest last
- Minimum of 48gb memory needed to start a 16gb LRM partition



Hints and Tips

- `uname -Ls`
 - Shows: 1 lparname
 - 1 = partition number
 - lparname = partition name
 - See Redpiece to get more information
- Resource Allocation
 - Desired
 - Minimum
 - Keep to bare minimum
 - Maximum
 - Set as high as possible
 - Applies to cpus, memory and I/O devices
- Cannot use graphics console to install
- Consider configuring empty slots into a partition



Hints and Tips

- Which LPAR is your service LPAR?
- How will you do installs
 - Allocate cd?
 - NIM?
- Backup Methodology?
- Create a partition layout in advance
 - Include devices, etc
- I/O devices are allocated at the slot level
- Boot disks – I/O drawer or 2104?
- 32bit kernel versus 64bit kernel
 - 32 bit supports up to 96gb memory
 - Need 64bit kernel to have more than 96gb in an LPAR
 - Need 64bit kernel for more than 16 processors in an LPAR



Configuration Info

5 lpars - 3 x 4way cpus, 2 x 8way, 3 x I/O drawers

| Drawer | Disks | scsi | Gb Fiber | Gb Ether |
|--------|-------|------|----------|----------|
| 1 | 4 | 2 | 5 | 2 |
| 2 | 4 | 2 | 5 | 2 |
| 3 | 2 | 1 | 5 | 1 |
| Totals | 10 | 5 | 15 | 5 |



Partition Map

| Lpar | Disk Drw 2 disks | CPUs | scsi Drw | scsi # | Gb Fiber Drw | Gb Fiber # | Gb Ether Drw | Gb Ether # |
|------|---------------------|------|----------|--------|--------------|------------|--------------|------------|
| 1 | 1 | 4 | 1 | 1 | 2,3 | 2,1 | 1 | 1 |
| 2 | 2 | 4 | 2 | 1 | 1,3 | 1,2 | 2 | 1 |
| 3 | 3 | 4 | 3 | 1 | 1,2 | 2,1 | 3 | 1 |
| 4 | 1 | 8 | 1 | 1 | 1,3 | 2,1 | 1 | 1 |
| 5 | 2 | 8 | 2 | 1 | 2,3 | 2,1 | 2 | 1 |



LPAR Notes – p650

➤ P650

- Internal disks and CD/DVD are on same scsi controller and assigned together
- Above can be moved using DLPAR
- Can order split backplane to split disks into 2 x 2 and then attach 2nd 2 to a separate scsi controller
- Boot disks
 - Use internal or disks in I/O drawer or an external 2104 disk drawer



LPAR Notes – p630

- Following must be assigned to a single partition as a group
 - PCI slots 1 and 2
 - Internal Ethernet 2 (U0.1-P1/E2)
 - Internal SCSI (U0.1-P2/Z1)
 - ISA based I/O (serial, keyboard, mouse)
- Dynamic LPAR cannot be used for those devices
- Parallel ports on a p630 are not supported in partitioned mode



References

- IBM Redbooks
 - The Complete Partitioning Guide for IBM eServer pSeries Servers
 - Configuring p690 in an IBM eServer Cluster 1600
 - pSeries – LPAR Planning Redpiece
 - Logical Partition Security in the IBM eServer pSeries 690
 - LPAR for Decision Makers
 - IBM eServer Certification Study Guide – p690 Technical Support
 - IBM eServer pSeries p670 and p690 System Handbook
 - Effective System Management Using the IBM HMC for pSeries



Questions

