Issue Date: AIX EXTRA eNewsletter
March 2004, Posted On: 3/1/2004

AIX EXTRA: Planning for LPAR

Jaqui Lynch

The introduction of logical partitioning (LPAR) and dynamic logical partitioning (DLPAR) with AIX V5 has caused a major change in the thought processes involved in planning in the pSeries world. Previously, correctly sizing a system involved pre-planning for resource consumption in terms of memory, CPU, DASD and adapters. While new dynamic partitioning features on pSeries partitionable systems have introduced flexibility in resource allocation that minimizes these constraints, other types of planning are now required.

LPAR enables workload consolidation by running multiple partitions on the same box. These partitions can be either Linux or AIX partitions, and both OSs can be run concurrently on the same server. However, there are some important items to remember when planning for this environment. The number of partitions available depends on the model and the number of processors installed. Some features such as Capacity Upgrade on Demand (CUoD) are only available on certain systems. (Note: pSeries Facts and Features, G320-9878, provides feature codes for partitionable hardware series and those machines that allow for CUoD.)

The first major change when moving to a system that will be partitioned (i.e., currently p630, p650, p655, p670 or p690) is the requirement of a Hardware Management Console (HMC). AIX V5 and the HMC are required for all partitioned systems. An 8- or 16-port asynchronous adapter can be installed in the HMC, allowing the HMC to control multiple pSeries servers with many partitions. The HMC is used to create and store partition profiles, start and stop partitions, display status, and provide a virtual console to connect to a partition. For non-partitioned systems, the HMC shows the LED codes. The HMC is required for management functions, but in the event of an HMC failure event, the managed partitions will still remain operational. The HMC must be backed up regularly to DVD, and these DVDs should be added to any offsite backup plan.

Currently, in a partitioned environment, each partition must have exclusive access to its assigned resources. The minimum assigned resources for an LPAR include one processor, 256 MB of memory, one boot disk, an adapter to access the disk, an Ethernet adapter to access the HMC and methodologies to run diagnostics (one LPAR must be assigned as the service LPAR) and for installation (e.g., NIM for installs). Resources are assigned at the slot level, which means that all disks attached to an adapter are assigned as one group to one LPAR. There are also specific rules with respect to assigning devices in the media drawer. (Note: Slot placement should be in accordance with the PCI Replacement Guide.)

Resources can be assigned to an LPAR as part of the minimum, desired or maximum requirements (i.e., it's possible to assign one CPU as a minimum, three for desired and eight for maximum). This means the partition will not boot unless it can get one CPU, it will always try to boot with three and the ability is there to increase the number allocated to up to eight using DLPAR. DLPAR provides the ability to dynamically add, remove or redirect resources to/from partitions. AIX 5.2 introduces this feature for processors, memory or PCI adapters, provided the HMC is at the correct level (post 10/2002 at least R3v2.2). This also enables the use of Memory and Processor CUoD, which allows currently unassigned processors and/or memory to be turned on and off. The added capacity can be either temporary or permanent.

When planning for memory, it must be noted that some of the memory is used by the pSeries to retain critical information. The amount that must be reserved depends on a number of factors, including the number of I/O drawers, number of partitions and how maximum memory for those partitions is set. At the bottom of real memory, 256 MB is reserved for the Hypervisor. At the top of memory, the Translation Control Entry (TCE) reserves memory to be used to translate I/O addresses to system memory addresses. This ranges from 256 MB to 1024 MB (in 256-MB increments) depending on the pSeries model and number of I/O drawers. Additionally, page table entries are required for each partition. The minimum reserved for page table entries is 256 MB, and each partition reserves 1/64th of the maximum memory value set in the HMC for its use (rounding up to nearest power of two). The page table entries for an individual LPAR must be in contiguous memory. On a p690, it's not uncommon for the overhead to be 2 GB with only two partitions. There are additional overheads associated if the OS is lower than AIX 5.2.

As you can see, planning is now even more critical. To get the best value from pSeries partitions, it's important to implement a resource plan and maintain planning worksheets and checklists to help ensure that resources aren't over-committed and that unnecessary overhead is avoided.