close window

IBM Systems MAGAZINE

Administrator

Print

Dynamic LPAR - The Way to the Future

June | July 2007 | by Jaqui Lynch

The introduction of LPAR and dynamic LPAR (DLPAR) with AIX* V5.2 caused a change in the thought processes involved in planning in the IBM* System p* world. New DLPAR features on System p platforms have introduced flexibility in resource-allocation constraints, but these changes require understanding and planning to be used successfully.

DLPAR is the facility on the IBM POWER4* and POWER5* hardware that allows users to dynamically move PCI slots, memory and processors between LPARs, increasing and decreasing resources on the fly. DLPAR is commonly used to move CD drives, DVD-RAM drives and tape drives between LPARs, rather than attaching one to each LPAR. Since its initial introduction with POWER4 and AIX V5.2, DLPAR has enhanced its capabilities and plays well within the new shared world enabled by POWER5 and AIX V5.3. DLPAR requires a hardware management console (HMC) at certain minimum levels and also requires that the OS be at certain levels - these can all be checked at the IBM HMC code Web site (see "References").

The Evolution of DLPAR

Since the AIX V5.2 environment is the basis for DLPAR I'll discuss that first. In an AIX V5.2 and/or POWER4 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 as everything on that adapter goes along with the device being assigned. In this environment DLPAR can be used to add, move or delete physical processors, memory and physical adapter slots.

In the AIX V5.3 and POWER5 (combined) environment, DLPAR can be used in the aforementioned manner, but it now offers some additional options. Processors in the shared pool can be moved in increments of one-tenth and memory in increments of 16 MB at a time. When working on shared processors, additional items must be changed, such as the weight of this LPAR relative to others, LPAR capping or uncapping and virtual processors. When increasing processor units, it may be necessary to increase the virtual processor count to accommodate them. The opposite may be needed when removing processors. Use the lparstat - i command before and after a processor or memory change to compare before and after.

Additionally, it isn't just physical adapter slots that can be dynamically moved in the 5.3/POWER5 world. It's also possible to use DLPAR operations on virtual adapters such as a shared Ethernet adapter or a virtual SCSI adapter. This adds more flexibility to an LPAR design. It should be noted that the addition or removal of any of these resources must be done carefully and the IBM documented procedures for each type of resource change should be followed. As an example, the CD-ROM should be removed (rmdev) from the LPAR before it's added to another LPAR. Certain actions will require you to run cfgmgr, etc.

The current versions of both Red Hat and Novell SUSE Linux* that run on POWER5 hardware are also enabled for DLPAR operations. The one exception is the moving or removal of memory as this isn't supported by the current (2.6) kernel.

More DLPAR

As with all DLPAR operations, the partition profile isn't changed when the resources are added or removed. Thus, if the profile isn't updated and the LPAR is shut down and reactivated, the resources must be added or removed again. If you want the action to be permanent, then it's important to also update the profile. DLPAR actions are constrained by two settings on the system for memory and processors - the minimum and the maximum. DLPAR can't add resources beyond the maximum and it can't reduce resources below the minimum. In the case of adapter slots DLPAR can't remove a required slot from an LPAR.

Only one DLPAR operation can be executed at a time and it can only act on one resource. Memory, processor or adapter slots must each be moved individually and sequentially; they can't be done concurrently. For the operation to work, the resource manager (IBM.DRM) subsystem must be running on the LPAR. If you get an error and think this might not be running, check it with one of the following commands:

```
lssrc - ls IBM.DRM
ps - ef | grep DRMd
```

Overhead and CUoD

The overhead for DLPAR operations is fairly minimal; DLPAR operations normally complete within a couple of minutes. The exception is memory, which is affected by the amount of memory being moved. According to IBM documentation, it takes between one and two minutes to remove 4 GB of memory depending on the state of the memory in the LPAR. One reason for this is the system must preserve the contents of the memory being removed - so it gets written to another location or copied to paging space. Additionally, filesystem buffer pools shrink when memory is removed. Applications may or may not recognize the fact that memory has been added or removed. This depends on whether the application is DLPAR-safe, -aware or -oblivious. By "oblivious" I mean that the application has no idea that any change has occurred. This is rare, but can have unforeseen results when memory or processors get removed.

Most applications today are DLPAR-safe. This means they remain unaffected by a DLPAR operation. They don't fail, but their performance may suffer if you remove resources since they won't be able to scale accordingly. Conversely, they won't necessarily notice new resources that are added.

Applications that are DLPAR-aware actually take notice when a DLPAR operation occurs and will try to adjust their system-resources use to take advantage of the new or modified resources. DLPAR-aware applications are always DLPAR-safe. The current version of DB2* is an example of a DLPAR-aware application.

DLPAR is used for memory and processor Capacity Upgrade on Demand (CUoD), which allows currently unassigned processors and/or memory to be turned on and off. The added capacity can be temporary or permanent. DLPAR also works with CUoD to provide dynamic sparing for processors, where an unlicensed processor gets allocated automatically to an LPAR to replace one that has reached the threshold for errors. In this case, the system requests that AIX vary on the new processor and then vary off the ailing processor.

Scheduling and Scripts

DLPAR operations can be scheduled in several ways. One option is to use the HMC and set up the DLPAR operation as a scheduled activity. This can be set up to run once or be repeated and can be time- or load-based. Additionally, there's a set of DLPAR scripts available or you can use Partition Load Manager (PLM).

IBM has made a set of primarily Perl scripts that can be used to control DLPAR operations available on the Alphaworks* site. These scripts have been around since at least 2004 and were designed to enhance the usability of DLPAR. The scripts allow you to display active partition configurations, monitor CPU and memory loads on multiple partitions and perform automatic DLPAR operations based on time or load. There's excellent documentation on the Web site on how to use the scripts, which are being used at many customer sites worldwide.

PLM

PLM is part of the Advanced POWER* Virtualization (APV) product that's purchased to enable Micro-Partitioning* and other features. It can be purchased as a separate product and runs on AIX V5.2 or V5.3. Not currently supported on the Linux OS, PLM provides automated processor and memory-resource management across LPARs. The LPARs are called managed partitions and the client partitions run an agent that talks to the PLM management server. User-defined policies are set up on the management server to control how and when resources are moved. For more information on APV, see the article "Take Home the Trophy With APV".

PLM can be used to view resources across partitions, group partitions into sensible groupings and to control resources across groups. A PLM manager instance can only control resources on a managed server, but the manager doesn't have to reside on that same server. It's possible to run multiple PLM management instances on one AIX system. Information is available online (www.ibm.com).

Future Technologies

DLPAR is the foundation of many advanced technologies that can be used to ensure scalability and reliability of the resources in the System p platforms. Coupled with CUoD technology and the automation scripts or applications, it's possible to set up the servers to be close to self managing, which is something that was only a dream back in 2001.

Recommended Reading

For more information on dynamic LPAR, check out the following reading materials, from the IBM Redbooks site (www.redbooks.ibm.com):*

- SG24-7940 - Advanced Power Virtualization on IBM* p5* servers - Introduction and Basic Configuration
- SG24-5768 - Advanced Power Virtualization on IBM p5 servers - Architecture and Performance Considerations
- SG24-7349 - Virtualization and Clustering Best Practices
- Red piece 4194 - Advanced Power Virtualization Best Practices
- Red piece 4224 - APV VIOS Deployment Examples
- The Complete Partitioning Guide for IBM eServer* System p* Servers
- System p - LPAR Planning Redpiece
- Logical Partition Security in the IBM eServer pSeries* 690
- Technical Overview Redbooks for p520, p550 and p570
- SG24-7039 - Partitioning Implementation on p5 and OpenPower Servers

References

As referenced in the article, you can find more information at:

- www14.software.ibm.com/webapp/set2/sas/f/hmc/home.html
- www.alphaworks.ibm.com/tech/dlpar
- http://publib.boulder.ibm.com/infocenter/eserver/v1r3s/topic/iphbk/iphbkconfigureplm.htm

- http://publib.boulder.ibm.com/infocenter/eserver/v1r3s/topic/iphbk/iphbkpolicy_file_configuration.htm