Automatic Partition Resource Manager for System i and iSeries

Understand the benefit of automating resource movement

Discover how to tune your system with APRM

See how APRM simplifies sharing I/O devices

Nick Harris
Dave Schnee
Peter Mullin

Archived

Redbooks

ibm.com/redbooks

Archived

International Technical Support Organization

Automatic Partition Resource Manager for System i and iSeries

June 2007

Archived

SG24-7174-00

Note: Before using this information and the product it supports, read the information in "Notices" on page v.

Archived

First Edition (June 2007)

This edition applies to Version 5, Release 4, Modification 1 of Automatic Partition Resource Manager® (APRM®).

© Copyright International Business Machines Corporation 2007. All rights reserved.
Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Archived

© Copyright IBM Corp. 2007. All rights reserved.

Archived

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:
This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

Archived

© Copyright IBM Corp. 2007. All rights reserved.

Trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

Redbooks (logo) ®
eServer™
iSeries™
i5/OS®
xSeries®
AIX®
AS/400®
IBM®
OS/400®
Redbooks®
S/370™
ServerProven®
System i™
System i5™

The following terms are trademarks of other companies:

APRM is a trademark of Barsa Consulting Group, LLC, in the United States, other countries, or both.

Excel, Microsoft, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Archived

Preface

Logical Partitioning for OS/400® opened the door to significant operational advantages and cost savings for IBM® iSeries™ and System i™ clients. You can now manage production workloads from many iSeries systems on a single newer server with LPAR function. And, in a System i world, you can now consolidate additional operating system environments onto a single System i footprint.

This IBM Redbooks® publication discusses setting up a partitioned IBM iSeries or System i platform. Our primary focus is to make it possible to share system resources between and among partitions easily, flexibly, and safely.

The contents of this book include information and a practical example of how to use an iSeries and System i tool, Automatic Partition Resource Manager (APRM) from Barsa Consulting Group, to help transform a company's infrastructure into an on demand, self-healing environment.

A complete copy of the APRM technical manual, Technical Alerts, and FAQs are available for download at:

http://www.barsaconsulting.com

There is a clear distinction in this book between iSeries and System i hardware. For the purpose of this book, iSeries refers to IBM eServer™ iSeries platforms with model numbers 2*xx*, 7*xx*, and 8*xx* that are capable of partitioning. System i refers to IBM System i hardware with model numbers 5*xx* and later models that are capable of partitioning.

Archived

The team that wrote this IBM Redbooks publication

This IBM Redbooks publication was produced by the International Technical Support Organization, Rochester Center, and Barsa Consulting Group, LLC.

Nick Harris is a Consulting IT Specialist for System i and iSeries in the Rochester Executive Briefing Center. He specializes in Availability, LPAR, System i hardware and software, external disk, and Integrated xSeries® Server for iSeries. He spent nine years in the International Technical Support Organization, Rochester Center, where he wrote IBM Redbooks publications and taught IBM classes worldwide. Nick spent 13 years in the United Kingdom (UK) AS/400® Business group and has experience in S/36, S/38, AS/400, and iSeries platforms.

© Copyright IBM Corp. 2007. All rights reserved.

You can contact him by sending e-mail to:

mailto:niharris@us.ibm.com

Dave Schnee is a staff consultant at Barsa Consulting Group, LLC, and is the chief architect of the APRM solution. Dave enjoys over 40 years of experience in designing and writing software, systems and business applications, specializing in the IBM mid-size systems market beginning with the System/3. Dave has a Master's degree in Electrical Engineering, advanced study in computer science and many IBM certifications.

You can contact him by sending e-mail to:

mailto:Dave@Schnee.com

Peter Mullin is a software solution practice leader at Barsa Consulting Group, LLC. Peter has broad practical experience introducing effective OS/400-based business solutions to almost all industries throughout the history of the platform. Peter's educational background includes a Master's degree in Business Administration.

You can contact him by sending e-mail to:

mailto:PMullin@barsaconsulting.com

Become a published author

Join us for a two- to six-week residency program! Help write an IBM Redbooks publication dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You'll team with IBM technical professionals, IBM Business Partners, and clients.

Your efforts will help increase product acceptance and customer satisfaction. As a bonus, you'll develop a network of contacts in IBM development labs and increase your productivity and marketability.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

Archived

We want our IBM Redbooks Publications to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review form found at:

  ibm.com/redbooks

- Send your comments in an e-mail to:

  redbook@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Archived

Archived

1

Executive summary

Archived

On Demand involves computers responding automatically to changing business conditions in self-tuning environments. Clients need self-running machines because many environments became complex and interwoven and the best way to keep them running is to replace human interaction. Studies show that the vast majority of downtime is caused by human error and tools to safely run repetitive tasks, provide continuous assessment, make decisions, and then self-heal without user input are highly valued.

Computing power today is significantly less expensive, per unit, than it once was, yet the costs are not so low that you can ignore them. There are many products that measure utilization of various parts of modern computer systems and attempt to provide management insights into possible adjustments and reassignment of resources to maximize the value realized from a given amount of computer hardware and software. Some of these solutions also have the ability to make many of these adjustments, subject to a company's business rules, all with the goal of providing better service at minimal cost.

© Copyright IBM Corp. 2007. All rights reserved.

Archived

2

Logical partitioning on the System i

Logical partitioning (LPAR) is a relatively recent technical development in iSeries computers. The LPAR allows one computer system to operate as two or more complete system images, sharing the hardware but otherwise operating somewhat or completely independently of each other.

Archived

© Copyright IBM Corp. 2007. All rights reserved.

2.1 iSeries LPAR History

IBM first introduced LPAR in 1988 on the S/370™ architecture. It was introduced in OS/400 in 1999 with V4R4. Operational advantages included centrally located hardware and software support and the ability to share hardware resources, such as processors, interactive Commercial Processing Workload (CPW), memory, terminals, tape drives, and CD drives. However, prior to V5R1, a big disadvantage was that resource allocation changes required an IPL restart.

Starting with OS/400 V5R1, IBM introduced dynamic LPAR, which allowed the movement of resources from one partition to another dynamically with no restarts required, provided that the allocation changes were guided by predefined minimums and maximums. Resource movements happened immediately but were initiated manually, scheduled, or embedded in program commands.

Figure 2-1 Requires OS/400 V5R1

In the V5R3 release, IBM eliminated the need for a primary partition to manage partition workloads, and IBM advanced to the Hardware Management Console (HMC). The HMC is a PC running a variant of the Linux® operating system with a command structure that is a fraternal twin to UNIX® commands. To ensure that no one can issue commands without authentication, the HMC requires that all foreign entry is through a “secure” protocol. Secure protocols include Secure Shell (SSH), Secure Copy (SCP), Secure FTP (SFTP), or one of the HMC specific clients, such as Web-Based System Manager (called WebSM), which is provided by IBM, or a Web browser using the HTTPS protocol.

On the System i5™, IBM further enhanced LPAR by providing *uncapped processors*. Uncapped processors allow a busy partition to temporarily borrow CPU resources from an idle partition, but when that idle partition gets busy again, the resource is automatically shifted back. Users cannot control this activity beyond established limits or by turning it on or off, so partition performance can be unpredictable. And what uncapped processors do with system resources over

Archived

time remains unknown, because there is no historical partition performance reporting.

2.2 LPAR advantages and disadvantages

The concept of logical partitioning, which allows one physical system to be split into two or more logical systems, is appealingly simple, but implementing a successful LPAR environment can be discouraging. You can gain many advantages by partitioning, but there are challenges as well. One challenge is that management of the system is somewhat more complex and might require a higher level of skill in computer operations and administration. Among the benefits is the ability to shift processing power to the workload that most needs it, which is not usually possible when the workloads are on separate physical systems.

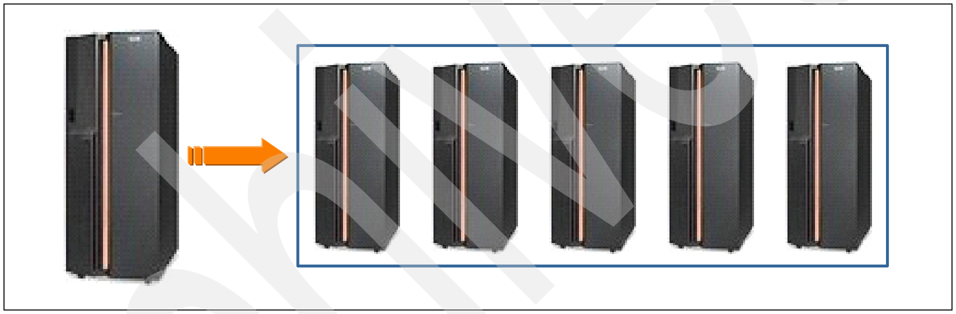


Figure 2-2 LPAR allows one physical system to be split into two or more logical systems

LPAR allows you to configure a single physical system as though it were two or more systems. Typically, you create each logical partition to perform specific tasks or to serve a distinct group within the organization. Each partition will have its own set of processing resource requirements. When you configure LPAR, you define the number of processors, the amount of memory, and the percentage of interactive CPW to assign to each partition.

A major challenge of logical partitions is to determine and then manage the proper allocation of processing resources among the partitions. Performance monitoring can reveal that one partition is starved for processing resources while another partition has processing resources to spare. Unfortunately, business operations are usually too complex to permit optimal resource distribution through manual or scheduled changes. Managing the configuration to ensure acceptable performance within every partition throughout the day can be overwhelming.

Archived

Businesses that require maximum performance from their partitions spend valuable time manually moving processing resources among partitions or, worse yet, masking the problem by overspending on processing resources.

Incorrect resource movement by untrained operators can cause catastrophic consequences. Studies show that human error causes the vast majority of system downtime, and the best way to avoid human error is to eliminate manual intervention. System tools that respond automatically to changing conditions and business needs are now available. Businesses are hunting for these autonomic tools because environments are becoming increasingly more complex and difficult to manage.

Consider that with V4R4 and V4R5, the movement of processing resources required a partition restart (IPL) to activate the changes. In addition, handling LPAR configuration changes required an experienced technician. This combination of finding time to IPL and the need for an experienced operator made regular LPAR changes impractical. V5R1 allows dynamic movement of processing resources without an IPL.

With V5R1 (and the appropriate version of hardware), IBM expanded LPAR capabilities to much smaller systems by using the ability to share one or more processors among multiple partitions. This ability to reallocate LPAR resources dynamically also enhanced the appeal of LPARs; with this capability, customers can use LPARs to manage varying workloads for different partitions throughout the workday or workweek. For example, a partition for East Coast locations can have more processing resources allocated early in the morning before business hours start on the West Coast, while a partition that is used by West Coast locations can have additional processing resources later in the day.

Shops that demand maximum processing power for each partition's workload spend too much valuable time manually balancing processing resources across partitions. As we have mentioned, managing the configuration to ensure acceptable performance within each partition throughout the day can be overwhelming. If the workloads in each partition do not commonly peak at the same times, it is possible to reallocate processing power among partitions to minimize processing time and make better use of the available hardware.

A daily operational challenge to the staff is the need to frequently back up data across multiple partitions. This challenge is met by reassigning the tape adapter slot and the attached tape drive between the partitions to be backed up. The devices, usually tape hardware, that can perform a backup quickly and reliably can be an expensive part of the total cost of the system. Again, if you can arrange the times of day to not overlap, you can timeshare a single, expensive, fast tape device among several partitions.

Archived

You can easily benefit from these LPAR advantages by using an automated tool that senses the workload and automatically adjusts the division of the total resources of the entire system among the partitions according to their instant workload and hardware needs.

Part of the solution to these challenges is a program product from Barsa Consulting Group named the Automatic Partition Resource Manager (APRM).

2.3 The APRM LPAR Management Tool

APRM is an iSeries and System i5 system tool that helps to transform a company's infrastructure into an on demand, self-healing environment. APRM automatically monitors and adjusts system resources among designated logical partitions. When you deploy APRM in a System i5 environment with or without uncapped processors, users can define business rules to tune LPAR performance. You can also tune memory, and you can access complete historical reporting of resource movements. Uncapped processors cannot move Input Output Processor (IOP) devices, such as tape and CD drives. Without APRM, untrained staff can easily move an IOP that controls disks, rendering a production partition useless.

APRM transcends the iSeries and System i5 standard adaptive processing functions because APRM governs the LPAR environment with predefined business policies and objectives. The APRM command interface safely runs repetitive tasks, eliminates the human factor, and provides the safety and confidence that clients are seeking. If clients do not use APRM, they need to predict resource needs and either schedule or move them manually. And that causes businesses pain.

APRM was the first product of its kind to reach the market (14 June 2002) and has six United States Patents for its technology. APRM, a ServerProven® solution, is the only LPAR management product that works with the HMC.

Three simple examples illustrate the operational advantages and cost savings that APRM delivers:

Example One A financial institution runs London operations in one partition and Los Angeles operations on another partition. When London goes to bed, Los Angeles starts up, and APRM automatically detects the workload shift and moves system resources from the London partition to handle the Los Angeles workload. The next day, APRM detects the shift to London and automatically moves resources appropriately. All of this occurs without manual intervention or inflexible, fixed scheduling.

Archived

Example Two The client needs to perform an application upgrade in a production partition. To avoid any production downtime, the client executes a planned role-swap from the production to another partition using high-availability software tools. But when production rolls to the standby partition, it is risky and time-consuming to require operators to also manually shift processing resources as part of that process. It is much better to embed APRM unique Command Language (CL) commands into the role-swap procedure commands to automate the process of system resource allocation among partitions.

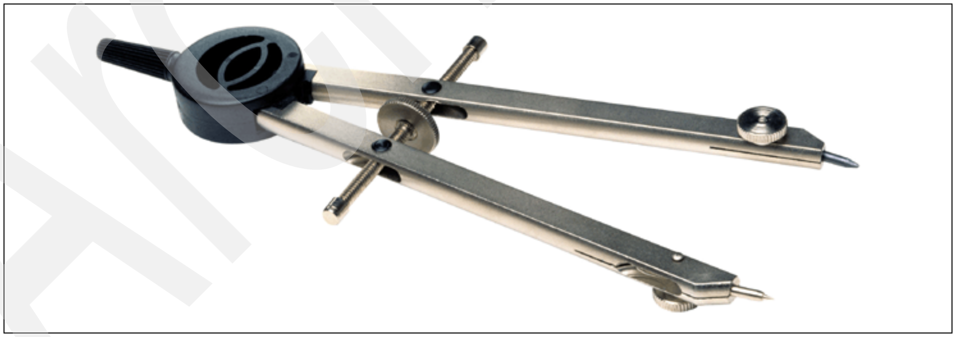
Example Three For smaller shops with less complex environments, making the most of existing hardware resources is key because it saves money. APRM eliminates the need to invest in redundant processing by maximizing limited system CPU, interactive CPW, and memory. APRM also allows all partitions to safely share a single tape or CD drive instead of investing in a backup device for each partition.

Archived

3

Setting up LPAR limits to make tuning possible

To improve performance and customize the operating environment, most clients in logical partitioning (LPAR) installations will modify standard system configurations.



Archived

© Copyright IBM Corp. 2007. All rights reserved.

3.1 General considerations

Task management can become extremely complex for businesses that use the iSeries system for multiple computing processes. Recent versions of the operating system (i5/OS® or OS/400) include the ability to consolidate systems through the use of LPAR, which allows you to configure one physical system to to operate as though it were two or more systems. In most cases, clients use logical partitions to define several smaller virtual systems within one large iSeries (System i) system. When you configure LPAR, you determine the number of processors, the amount of memory, and the amount of interactive Commercial Processing Workload (CPW) to assign to each partition. We refer to these three resources, taken together, as *processing resources*.

If you plan to implement an LPAR environment that includes dynamic resource sharing, answering “yes” to one or more of these questions might make you a candidate for dynamic LPAR (DLPAR):

- Are there common time or processing-period boundaries where resources can be moved between partitions to gain some advantage?
- Do certain partitions have peak CPU or interactive processing times?
- Do certain partitions have peak resource needs (for example, CPU or interactive CPW)?
- Are there off-peak processing times when a partition can relinquish resources?
- Do certain partitions have specific processing trends (for example, daily flash reports) that require a particular type of resource?

Before taking advantage of dynamic LPAR, it is a good idea to first analyze your current environment by documenting each partition’s system requirements, monitoring system activity, and then analyzing the results. Next, we describe brief descriptions of the steps that are required to analyze partition requirements. We provide this analysis to help you devise a plan for timely, dynamic movement of resources:

- Describe each partition’s critical jobs, jobstreams, and processing functions.
- Understand the processing windows needed for each partition.
- Gather performance data by using Collection Services. Gather data by partition, by day, by week, and by month.
- Consolidate and analyze data using a tool such as Performance Monitor to create an overall picture of system resources.

When you finish analyzing the initial processing requirements of each partition, partition configurations, and planned resource movement, it is imperative that

Archived

you document the process. Be sure that your documentation includes the requirements that were defined early in the process, the performance analysis summary, the partition configuration details, the resource movement guidelines, and the performance metrics to use for evaluation purposes. Your documentation also must include a plan for regular performance evaluation and a plan to update the process when necessary. Unfortunately, a haphazard plan can cause a dramatic reduction in overall system performance. The lack of a thorough evaluation causes eventual, if not immediate, performance problems, as well as potential systems management complications.

By implementing stringent requirements for gathering and planning, you can define an LPAR implementation plan that provides maximum performance and optimal use of resources with minimal administrative intervention. In a well-defined operating environment, resource utilization can improve dramatically.

Resource movement can levy a “tax” on your system. Before devising a plan for resource sharing, we examine several of these performance taxes. In general, regardless of the method that is used to move resources or the resource to be moved, it is important to avoid quick overreactions to a partition that needs immediate resources. Moving resources from one partition to another partition when the resource need is short-lived causes unnecessary churn as the system moves resources. This, in turn, causes a reduction in resource efficiency, the opposite goal of most LPAR users. Indiscriminate memory movement from one partition to another partition can affect performance.

Another small performance tax occurs in the re-creation of SQL access plans that have been stored in memory. Storing access plans in memory can force the system to rebuild those access plans in a partition where memory is being removed. As access plans are rebuilt, SQL response times can increase. Rebuilding the access plans usually takes less than 20 seconds.

You can accomplish dynamic resource sharing through manual intervention, through scheduled moves, or through the use of programs that determine a valid need for additional resources. After you understand the configuration constraints, allocating resources for an iSeries system partition through either System Service Tools (SST) or iSeries Navigator is relatively simple. iSeries Navigator provides a step-by-step wizard to create a new partition. On System i platforms, similar tools are available by using the Hardware Management Console (HMC).

Each partition must have a certain amount of interactive CPW and CPU assigned to it. As you increase the amount of interactive capacity, the system might require additional processing units, or dedicated processors, to support the assigned interactive percentage. LPAR configurations entail defining the amount of system resources assigned to each partition. Each partition has lower limits and upper limits for its processor and memory values. You set these limits

Archived

differently for iSeries and for System i hardware, but after you set them and the partition starts, you cannot change these limits without a partition restart (IPL). For iSeries systems, some of these changes require a full system restart. For System i platforms, there is also a recommended amount of these resources to consider.

Certain systems have a feature called *On Demand* that relates to processors and memory. The concept is that IBM builds a machine with more hardware than is initially activated, and the client can turn the capacity on as needed and not pay for the additional hardware until it is activated. Memory is sold by size (megabytes or gigabytes); processors are sold by number and speed (measured in gigahertz and CPW). Certain systems have limited but not zero interactive capability, which is sold by CPW but is not subject to On Demand activation. Other systems are sold for batch workload only (zero interactive capability) or as Enterprise Edition systems with no restriction on how you use the total workload capability.

For any dynamic LPAR operations to be possible, there must be some flexibility between the lower limits and upper limits of one or more of these processing resources. For this reason, it is usually advantageous to set the limits as wide as possible with one exception: The upper limit for memory in a partition controls the amount of system memory that gets dedicated to managing that memory on behalf of the partition. This is sometimes called the *hardware page table.* On a large machine with many partitions, the hardware page table can amount to gigabytes of memory if the upper limit for memory is set high in many partitions (especially where most of these maximum values cannot be realized simultaneously). Consider a system with 64 GB of total memory and eight partitions to be configured. There needs to be sufficient space in the hardware page table for at least the 64 GB of real memory. If, however, each of the eight partitions has a maximum memory limit of 40 GB, there must be hardware page table space for 320 GB (eight partitions times 40 GB for each partition), which takes about five times the minimum space needed. To make adjustments possible, the sum of the maximum memory for all the partitions should be larger than the total physical memory, but perhaps not this much larger. There is a trade-off here of flexibility as compared to cost. However, this hardware page table size is less important for System i platforms than for iSeries platforms because of the granularity of memory assignment. On iSeries hardware, memory is assigned in units of 1 MB; on System i platforms, the granularity is between 16 MB and 256 MB, depending upon the total amount of installed memory. It is also possible, by user action, to force a larger granularity than the default for the installed memory. This larger granularity means that there are fewer entries in the hardware page table for each GB of assignable memory.

Removal of memory from an active partition takes time, sometimes more than a few minutes. With OS/400 or i5/OS, the memory to be removed must first be

Archived

moved from whichever pool it is assigned, through the machine pool, then into the base pool from which it can be removed. The content of the memory must be saved to DASD (if it has not already been saved), and the memory must be cleared to prevent the possibility of data from one partition being read by another partition. Much of the time required for this operation is due to the time needed to write the content to a Direct Access Storage Device (DASD). In addition, the system value QBASPOOL sets a minimum amount of memory that must remain in the base pool, and if this value is set too high, the memory is not relinquished by the partition. This causes the DLPAR removal of memory to fail to a certain extent.

There are two ways to configure processor power for a partition: with either dedicated processors or shared processors. With dedicated processors, the partition has an integral number of real physical processors assigned to it. With shared processors, a number of processors are assigned to the processor pool, and this pool of processor power is assigned to one or more partitions in units of 0.01 (equivalent) processor.

In reality, there is no such thing as a fraction of a processor. What really happens is that a partition is assigned one or more real processors for some fraction of the time, and those same processors service other partitions for the rest of the time. The switching among partitions is rapid and frequent enough that it appears that the computing power is shared in proportion to the assigned processing units. There is, of course, some inefficiency with this method of sharing, because each processor must be interrupted periodically, its state saved somehow, and the prior state of some other partition’s job restored before the processor can resume effectively servicing that other job. Yet, this method of sharing offers great flexibility in putting the processing power where you need it.

There is also the concept of *virtual processors*, which is another value that you must set with lower limits and upper limits. This is not a measure of how many processors to assign, but instead it is a measure of the maximum number of tasks that the operating system attempts to dispatch simultaneously. It is usually inefficient to set the number of virtual processors higher than the number of actual processors assigned (or the number of shared processing units rounded up to an integer). However, on System i platforms, the concept of uncapped processors, which we describe later in this chapter, might modify this general rule.

On many iSeries platforms and some of the System i platforms, there exists an assignable amount of interactive performance percentage that you can, and sometimes must, assign. Many of the larger System i platforms do not allow the specific assignment of interactive percentage; these systems have either full or zero interactive capacity. Where you can assign interactive performance, this metric also has lower limits and upper limits, generally between zero and one

Archived

hundred percent, and often must be assigned in a measure that depends upon the amount of total processor power that is assigned to the partition. For certain System i platforms, you must also consider the desired amount of interactive performance percentage.

Where the interactive power of the system is a large proportion or all of the total power, there is a rule that the interactive percentage assigned cannot exceed the total power that the processors assigned to that partition can provide. This is clearly shown in the case of an Enterprise Edition system where interactive performance is unrestricted, and only certain processors are activated (using Processor On Demand technology). For example, on a system with 16 installed processors but only nine activated (56.25% of the installed processing power) and full interactive capacity, it is impossible to assign 100% of the interactive power because those nine processors cannot service 16 processors' worth of interactive load.

Interactive workload is a concept that applies only to OS/400 and i5/OS partitions. On iSeries and System i hardware, you can configure partitions to use Linux for an operating system, and System i platforms can support AIX® as well. You can assign these partitions an interactive performance percentage, but they do not use it, and if interactive capacity is limited, this is simply wasteful.

3.2 iSeries systems

The first consideration for iSeries systems is determining what function the primary partition plays in the overall system-use plan. Although on the surface, all partitions seem to be created equal, the primary partition has special responsibilities. It is responsible for all partition management functions, and it must be active before any other partition can be available for use. Restarting the primary partition also restarts all secondary partitions. This requirement increases the importance of thinking about how you use the primary partition. We do not advise that you use the primary partition as a test environment, for example, if your test environments tend to require frequent, and often unplanned, system restarts. We advise that you create a small primary partition that is strictly used to manage the overall system, relegating other functions and activities to the secondary partitions.

Next, consider how you will allocate processors to the partitions. You have two options: They can be distributed as shared processors (fractional increments) or as dedicated processors (whole). Both methods have advantages and disadvantages. Shared processors offer greater granularity in the distribution of processing power because you can distribute processors in increments as small as one-hundredth of a processor. You can configure a system with one or two processors to act as several logical systems, which offers better overall use of

Archived

available processing units. The downside is that shared processors increase the system overhead because of the maintenance of the processor cache.

The concept of shared processors does not mean that the system physically splits the processor among partitions. Actually, each partition uses 100% of the processor for its allocated time slice. For example, a partition that is assigned one-quarter of a processor has full use of the processor 25% of the time. The process of switching from one partition to the next forces the processor cache to be cleared, which, in turn, can cause as much as a 10% overall system performance decrease. You must assign at least one-tenth of a processor to any partition that uses shared processors. Although it is possible to allocate such a small portion, IBM recommends that you assign a minimum of one-quarter of a processor to each partition.

Dedicated processors offer the converse, that is, more efficient use of processing power. Again, the performance gain is due to the absence of the need to frequently clear the processor cache. Yet, the use of physical processors can leave unused CPU in one partition while another partition experiences high CPU usage. Additionally, it is often difficult to spare an entire processor for a partition that uses large amounts of CPU.

A component (the hypervisor) manages all of the LPAR activity. The hypervisor runs as part of the System Licensed Internal Code (SLIC) and takes resources assigned to the primary partition. For that reason, there is a larger minimum memory requirement and more processor workload in that partition than necessary for just user tasks and normal operating system tasks. You can assign all of the installed memory (and activated memory if On Demand memory is installed) to the partitions. But if your partition definitions of maximum memory allow all of the memory available on the system for every partition, not only is the page table excessively large, it is physically impossible for all partitions to use the defined maximum memory at the same time. This illustrates the importance of your understanding the requirements of each partition and of not overestimating the maximum amounts of resource. In addition, redefining a partition’s minimum values and maximum values requires a restart of the partition, so it is better to define the range accurately from the beginning. As you configure processor values and interactive minimum values and interactive maximum values, you see that their values are tightly integrated.

You can assign resources from the primary partition only. There are traditional 5250 application interfaces: SST and Dedicated Service Tools (DST). There are also iSeries Navigator and a single OS/400 Application Programming Interface (API) named QYHCHCOP.

As you configure the partitions, you must set minimum values and maximum values for memory, processors, and interactive percentage. Setting these values is not as easy as you might think. All hardware resource assignments are

Archived

maintained in the primary partition. The page table size is based on the potential need for memory if all partitions operate at their maximum settings.

You assign processor and interactive power to each partition. They stay assigned that way until you change them manually or they are changed by a program that invokes the QYHCHCOP API. A partition that is not started keeps all of its resources (in effect, wasting that processing power) and can be started at any time with its resources unchanged from the last time it was running.

If the actual assigned memory does get changed while the partition is powered off and this is an i5/OS (OS/400) partition, it is good to have the system value QPFRADJ set to a value of 2. Setting QPFRADJ to a value of 2 allows the adjustment of the QBASPOOL system value and allows the memory assigned to shared pools to be changed automatically at IPL and during normal operation. Conversely, if no adjustment is allowed and some memory was removed while the partition was powered down, the partition might be able to start but OS/400 might not.

Interactive power is assigned to partitions in increments of 1%. At V5R1 and earlier, the minimum value is 1 (one) percent; at V5R2 and later, the minimum value possible to assign is 0 (zero).

In general, assigning the lowest memory allowed by SST, DST, or iSeries Navigator is likely to cause operational problems in an OS/400 partition. There will be insufficient memory for any significant Reserved memory (a column shown on the WRKSYSSTS panel), certain system functions might not be able to start or run, and performance is likely to be marginal at best. When you upgrade the operating system, the partition might be unable to IPL at all, or in some cases, the hypervisor might reassign all of the system resources to the primary partition, and a manual rebuild of the LPAR configuration might be necessary.

Archived

3.3 System i platforms

It is helpful to discuss the basic changes to the newer System i hardware. In June of 2004, IBM introduced enhancements to iSeries systems that included software, V5R3 i5/OS, and hardware called System i platforms. IBM identifies these newer models as any 5*XX* model number.

Fundamental differences in managing LPAR on the new System i platforms result from two key changes to the LPAR environment. The first change is the elimination of the primary partition. With the new System i hardware, all partitions are created equal. You no longer designate one partition to manage LPAR functions, such as assigning processing resources for each partition, or sharing and moving hardware, such as tape drives used for backup and recovery processes.

Closely related to the elimination of the primary partition is the second LPAR change, the introduction of the HMC. The HMC is a single-purpose closed appliance for controlling multiple System p or System i platforms. An HMC connects to the Service Processor (SP) of each POWER5 server through an Ethernet network and allows the servers to be managed, including power up and power down, starting, stopping, and configuring Logical Partitions (LPARs), and installing firmware. An HMC is an IBM PC with special software (referred to as machine code) installed. A systems administrator can sit at the HMC or remotely access it. An advantage of the HMC and its interface with the SP is the ability to dispense with the need to restart (IPL) partitions, especially the primary partition, which effectively means the IPL of all partitions, regardless of the system-level configuration changes. For example, if you add a new partition to an iSeries system, you must IPL the entire system. Alternatively, with the new System i platform, you can create new partitions without interruption to existing partitions.

The Service Processor (SP or FSP for Flexible Service Processor) controls all of the LPAR activity, and the attached HMC provides a user interface for this control. The FSP has its own processor but takes a certain amount of the installed main memory for its own use and for the hardware page table. The assignment of this memory to the FSP is in increments of the same granularity as other memory allocations, and you cannot assign this *Firmware memory* to any active partition. As you activate more partitions, the Firmware memory requirement usually rises. This can be confusing when you look at the available memory and then try to activate a partition that needs all of it. The activation of that partition can fail due to insufficient memory because the Firmware memory first takes what it needs from the available memory and then attempts to allocate memory to the partition.

Processor power, interactive power (if it is assignable), and memory can change while a partition is inactive. For processor power, interactive power, and memory, you assign a lower limit value, a desired amount value, and an upper limit value in a partition profile. You can assign several profiles to each defined partition, and you can mark one of these profiles as the default profile to use in the absence of a specific designation for a single activation. The activation of another partition or a DLPAR operation can remove part or all of the resource formerly used by a currently inactive partition. Unless the minimum resources listed in the profile are either still assigned to the partition or available, you cannot activate that partition. In some fashion, this allows you, when managing the system, to make more complete use of system resources by not forcing them to stay idle when one or more partitions are inactive.

Another feature of some System i platforms is the concept of *uncapped processors*. This concept applies primarily to partitions that are configured to use processor power from the shared processor pool. Uncapped processors allow processor power that is unused by the partition to which it is assigned to be "borrowed" by another partition that is already using all of its assigned amount of processor power. Because this borrowing occurs on a fraction-of-a-second basis, it can be a great help to a system that can benefit from more raw computing power without any increase in memory, interactive performance, or DASD speed.

There is also a parameter called *uncapped weight*, which serves to control the allocation of unused processing cycles among several partitions wanting to borrow processor power. If a partition uses several times the processor power that it is assigned, the WRKSYSSTS display shows several hundred percent of CPU usage.

If there is little-to-no unused processor power in the processor pool, uncapped processors do not improve the performance of the system because there is no otherwise wasted processor power to reassign. Because the borrowing of processor power can cause the real number of processors that are used by a partition to be higher than the number actually assigned, it can be advantageous to allow the dispatching of as many tasks simultaneously as there are processors in the processor pool.

There is usually a restriction to the maximum number of these virtual processors according to the current actual processing units that are assigned. In many systems, there can be at most 10 virtual processors for each full real processor or one virtual processor for each 0.10 processing units that are assigned. This can reduce the assignable number of virtual processors in certain cases. For example, consider a partition with 0.25 processor units assigned. For this partition, there are a maximum of two virtual processors allowed. If the processor pool has eight real processors assigned and no other partition is using any significant processor power, this partition can still utilize only two processors at any time.

The uncapped processor designations for partitions using the shared processor pool are different from those using dedicated processors. For partitions that use the shared processor pool, the values allow the partition either to use (borrow) processor power or to prohibit this action. For partitions that use dedicated processors, the values work in the opposite manner; they either allow the partition to share its own processors when they are idle or they prohibit this action.

Archived

4

Setting up APRM limits to make tuning efficient

Automatic Partition Resource Manager (APRM) brings harmony to logical partition operating environments.

Archived

© Copyright IBM Corp. 2007. All rights reserved.

4.1 The APRM resource management tool

APRM is an iSeries system application that dynamically monitors and adjusts processing resources among designated partitions. Resource movement occurs when APRM detects high levels of "resource needs" within any partition and, at the same time, detects the availability of processing resources in other partitions.

The benefits of APRM include:

- Automatic, continuous balancing of processing resources to maximize iSeries and System i performance
- Quick, simple configuration of tuning rules to customize your APRM environment
- Required processing resource movement without operator intervention
- Efficient use of processing resources among partitions
- Constant monitoring and reporting of partition workloads
- Straightforward command interface to switch Input Output Processors (IOPs) between partitions
- Historical view of processing resource demand and movement available for analysis

APRM brings harmony to logical partition operating environments by automatically moving processing resources to those partitions that show distress in any of three areas: CPU utilization, interactive workload, or memory page faults.

Regardless of the model of iSeries or System i, you must install APRM on all partitions in which automatic tuning, monitoring, or reporting of partition activity occurs. In addition, you must have APRM actively running in any partition that participates in the automatic movement of an IOP from one partition to another (that is, the requesting partition or the partition using the IOP). If you use an iSeries system, you must install and start APRM on the primary partition; if you use a System i platform, you must install and start APRM on every partition that participates in tuning, performance reporting, and automatic IOP switching.

APRM works to adjust processing resources according to the need. APRM adjusts processing resources only within lower limits and upper limits separate from those that are used by the LPAR hypervisor or Service Processor (SP). Unlike those system LPAR limits that you can only change with a partition restart, you can change APRM tuning limits (rules) at any time, and these changes are effective immediately. The APRM limit rules must be within the IBM limits, because you cannot violate the system LPAR limits for a partition.

Archived

The APRM limit rules consist of lower values and upper values for processor power (either dedicated processors or processing units, depending upon the processor type for each partition), interactive power (if it is assignable), and memory. APRM calculates the appropriate value for virtual processors automatically. There are also “Yes/No” questions for whether to implement tuning and whether to respect the IBM LPAR minimum values or the APRM (possibly higher) minimum values when the partition is inactive or powered off.

On System i hardware, APRM establishes a Secure Shell (SSH) connection from each i5/OS partition to each Hardware Management Console (HMC) controlling the system on which they run. This connection is a two-way conduit that mimics the operation of the HMC command-line interface. Through this channel, APRM can derive information about the LPAR configuration and perform dynamic LPAR (DLPAR) operations to make resource changes.

There is also a priority for tuning. This is a number between 1 and 99 where, in a manner similar to OS/400 job priority, the lower number is the higher priority. The priority of APRM, however, is proportional rather than absolute. That is, a partition whose priority is 40 is allowed twice the effective workload slowdown than a partition whose priority is 20. The definition of effective slowdown (“pain”) is more complex than simply processor utilization percent or page faults per second. It takes many factors into account, combines interactive usage with processor usage (where interactive usage is a factor), and tempers its usage of page fault rates with the number of Direct Access Storage Devices (DASD) devices and the CPU utilization. This measurement and evaluation are the subject of one of the six patents awarded to the APRM development team.

DLPAR operations that modify either the amount of processing resources and interactive performance percentage are completed rapidly, but moving memory between partitions takes time, and the length of time depends on the utilization of the resources. The more memory that is taken from a partition, the longer it takes. Because of the time needed to remove memory from a working partition, it might not be advantageous to allow a large variation in memory levels during time periods when several partitions are heavily loaded. If, however, processor levels are adequate and the major resource shortage is likely to be memory, the memory is the resource to tune despite the time it might take.

Because APRM offers several ways to change tuning rules, both interactively (with the 5250 applications and iSeries Navigator plug-in) and through an OS/400 command, you can prearrange several sets of tuning rules and store them in simple Command Language (CL) programs with different minimum levels and maximum levels of resources for each partition appropriate to different cycles of workload. For example, consider a system with a partition dedicated to California, USA, and another dedicated to London, England. Each partition has a workload during its regular business day, which is followed by nightly batch and

Archived

backup work, whose response time is less critical than the response time during the business day. If memory is really rededicated twice each calendar day, then it can be reasonable to raise London's minimum memory setting significantly in advance of the start of London's business day (perhaps an hour or two) to ensure that the memory has moved to that partition before it is needed. Near the end of London's business day, its minimum memory value can be lowered to allow more memory to be moved. At that particular time of day, it is doubtful that much memory is needed elsewhere. Similarly, California can raise its minimum memory limit in advance of its needs and later lower its minimum memory limit in advance of London's needs.

APRM takes advantage of the flexibility of dynamic LPAR to move resources among partitions. However, APRM respects the limits that are set up in your basic LPAR configuration. To optimize the capacity of APRM to enhance your overall system performance, it is important to review your basic LPAR configuration with the following recommendations in mind:

- When using shared processors, do not assign less than 0.25 of a processor to any partition that will have any application workload. Although you can allocate as little as one-tenth (0.10) of a processor, IBM recommends this minimum allocation only in extremely limited conditions. Refer to IBM LPAR documentation for specifics, but note that newer processors are much faster than they were when this 0.25 processing unit rule was issued.
- Keep processing resource minimums low for all partitions to maximize the flexibility of dynamic LPAR.
- When setting minimums, consider the minimum amount of resources necessary to properly support all of the tasks associated with the partition start-up process.
- Keep processor and interactive capacity maximums high to maximize the flexibility of dynamic LPAR.
- Keep memory maximums reasonable. Needlessly large memory maximums require large page tables to keep track of memory assignments and waste system memory, although with System i memory granularity, this is not as wasteful as it is on iSeries hardware.

4.2 iSeries systems

Figure 4-1 on page 23 represents the basic architecture of APRM and the major components of an iSeries system. Remember that all APRM tasks are, at some point, completed through the primary partition. In this example, all partitions use OS/400.

Archived

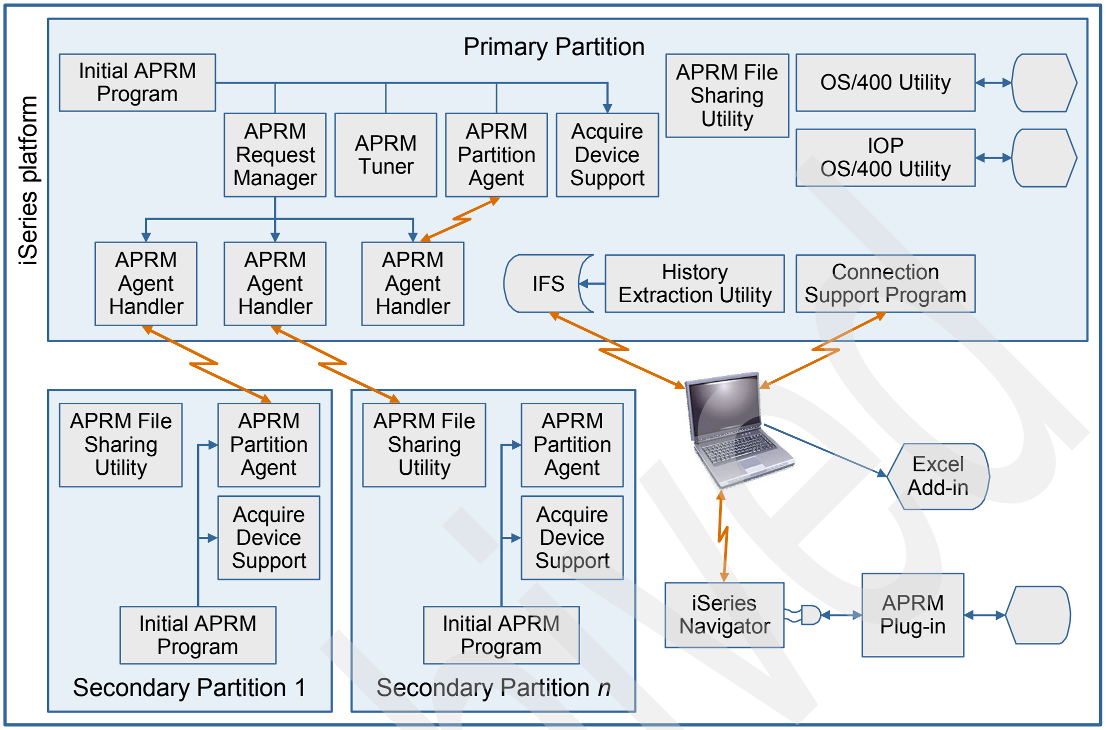


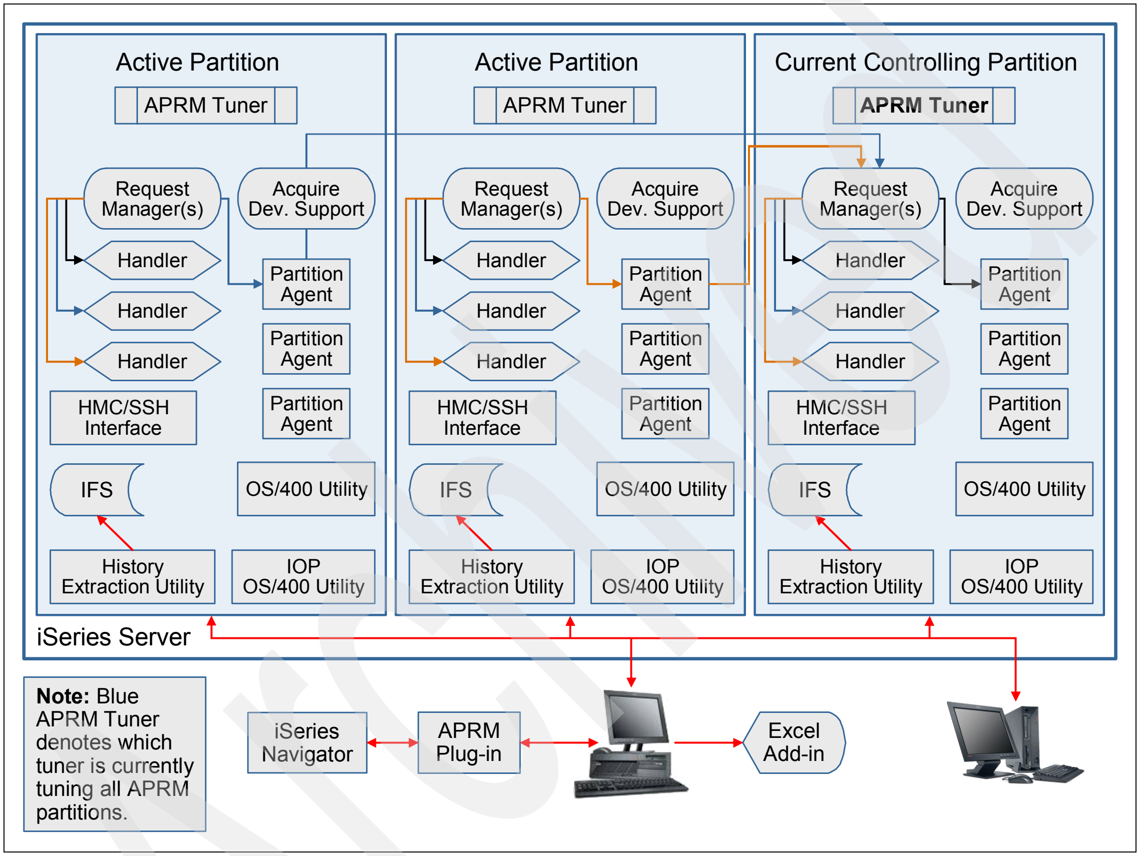
Figure 4-1 iSeries architecture

Because of the way that the iSeries LPAR hypervisor works, APRM can control the minimum values of resources for partitions that are powered down; DLPAR operations are permitted for all partitions whether or not they are active.

4.3 System i platforms

APRM establishes an SSH connection from each i5/OS partition to each HMC controlling the system on which they run. This connection is a two-way conduit that mimics the operation of the HMC command-line interface. Through this channel, APRM can derive information about the LPAR configuration and perform DLPAR operations to change resources.

Figure 4-2 on page 24 represents the basic architecture of APRM and the major components on a System i platform. In this environment, you can manage APRM from any of the three partitions shown. In Figure 4-2 on page 24, the partition on the far right currently manages all APRM tasks. Note how all three partitions contain all of the APRM programs. In the controlling APRM partition, the APRM tuner that is currently responsible for partition resource allocations and hardware-sharing tasks is the APRM tuner on the right. The order-of-succession list that you define during the configuration of APRM determines the APRM controlling partition.



Note: Blue APRM Tuner denotes which tuner is currently tuning all APRM partitions.

Figure 4-2 System i architecture

On iSeries systems, APRM completes all processing of resource tuning and hardware movement through APRM programs running on the primary partition. On a System i platform, you can potentially use any partition to manage LPAR performance. It is important to note that because there is no longer a requirement for one partition, known as the primary partition, to be active at all times, APRM has no guarantee that any one partition can always be used to manage APRM tasks and monitor partition activity.

Archived

To accommodate this change, you must now configure APRM on System i hardware with *partition succession* control. Partition succession in its simplest terms is a prioritized list of partitions that you define and that APRM uses to determine which partition to use to manage APRM tasks. If the first partition on the succession list becomes inactive for any reason, the next partition in the list becomes the active APRM partition with respect to managing all resource allocations and sharing hardware. APRM continues moving down the succession list as necessary. As partitions that are higher on the succession list again become active and available, APRM returns management control to that partition.

Because of how the HMC or SP allocates resources and the fact that DLPAR operations are not permitted to and from partitions that are inactive, APRM cannot guarantee that resources remain in any definite amount when a partition powers down. Therefore, APRM can only refuse to use resources elsewhere that might make each controlled partition's minimums impossible to meet, and then only if APRM controls all the partitions and their resource changes.

Simply activating an unrelated partition can cause another powered-down partition to lose resources. The HMC or SP attempts first to provide a newly starting partition with its minimum needs, then tries to satisfy its requested resource levels. While doing so, the HMC or SP removes resources as needed first from inactive partitions down to their most-recently used minimums, and then, if the HMC or SP still needs more resource, it completely robs an inactive partition. In this case, either a DLPAR operation is needed or another partition must be powered down before the robbed partition can be activated.

Archived

Archived

A real-life practical example of LPAR configuration

In this chapter, we show you hardware configuration and partition requirements from a real-life situation.

Archived

© Copyright IBM Corp. 2007. All rights reserved.

5.1 The hardware

In the following example, the system was a medium-sized iSeries model 870, but except for the primary partition requirement, it is comparable to a System i model 570. This was an Enterprise-class machine; that is, there was no limit on the interactive workload other than the limit imposed by the number of processing units themselves. We show both the (real) iSeries configuration and its (hypothetical) replacement System i configuration.

5.1.1 Processors

The iSeries was a Processor On Demand model. There were 16 installed processors, but only nine were active. This allowed the client to purchase the machine for several hundred thousand dollars less than a machine with all of the processors active but provided for future growth with no disruption of his business. The hypothetical System i replacement was a 570 with eight processors but with only five activated processors. The total Commercial Processing Workload (CPW) of this System i configuration exceeds that of the model 870 that it replaces.

5.1.2 Memory

The total installed memory was 48 GB; no Memory On Demand was installed, so all 48 GB were available for use. The theoretical System i replacement we discuss has 64 GB for improved performance and growth.

5.1.3 Interactive workload

Despite the fact that this was an Enterprise-class machine, but because it was an iSeries model, we considered the interactive workload allocation. Because it was an Enterprise-class machine, the interactive percentage was always made to track the processor allocation whenever a change to the actual configuration was made. To do anything else makes no sense, because the remaining interactive resource cannot be assigned to any other partition because it is impossible to make the interactive workload greater than the batch workload limit. All of the partitions on this machine used OS/400 (i5/OS); none of them used Linux.

Because only nine of the installed 16 processors were active, the client could allocate only 57% of the total machine's interactive capacity (9/16 = 56.25%). For an iSeries machine, allocation of interactive capacity is in whole percentages, so IBM allows you to round up the total. On machines with all processors activated (or no Processor On Demand capability), you can allocate 100% of the interactive capacity.

Archived

On Enterprise class System i hardware, there is no requirement nor any ability to allocate interactive workload; you can assign only processor power, virtual processors, and memory. In addition, for any System i platform, instead of setting an exact current value for each metric, you can specify a desired value, which will be allocated if it is available when you activate the partition by starting it or performing an IPL. However, remember that the partition might be started with only the minimum of any or all of these metrics if more resources are unavailable. If this were a much smaller System i platform with limited, but not zero, interactive capacity, the allocation of interactive capacity is specified in units of hundredths of a percent.

5.2 The challenge and the LPAR requirements

The business needs dictated eight client-usable partitions that we describe in Table 5-1. Because this was an iSeries system, a “skinny” primary partition was also configured. If this had been a System i platform, there is no primary partition, but instead a block of memory is earmarked for the Service Processor (SP). We refer to this block of memory as *Firmware memory*, but there is no processor power required for the SP.

Table 5-1 shows the partitions and their intended usage.

Archived

Table 5-1 The partitions and their intended usage

Partition	Intended use
Partition #0	The “skinny” primary partition. This partition’s purpose is solely to control logical partitioning (LPAR) and I/O for the rest of the system. It does not exist in the System i data.
Partition #1	A full business enterprise system for the company’s largest division. This partition will try to use at least half of the processing power of the whole machine for a significant fraction of the day and night.
Partition #2	A full business enterprise system for the company’s second-largest division. The workload is between one-half and one-third that of Partition #1. The worst time for contention is in the 5 a.m. to 7 a.m. time range on most days.
Partition #3	A business partition with modest processing requirements; the workload seldom will need more than one-half of a single processor.
Partition #4	A business partition with modest processing requirements; the workload seldom will need more than a single processor. The management of the company deems this partition less important than the other business workloads.
Partition #5	A development and testing partition for a new line of business. This partition will have light usage and then only sporadically. A fraction of one processor will suffice until the business unit is brought online. Until then, this partition is considerably less important than other business workloads.
Partition #6	A working partition for a small line of business. This partition will have only light usage and then only sporadically. Less than one processor will suffice most of the time, but there will be sporadic bursts of more significant activity.
Partition #7	A development partition solely for the program developers. For much of a typical 24-hour day, the partition will be idle, but there will be times when two or three processors can be busy for more than an hour.
Partition #8	A testing partition for a soon-to-be-implemented Data Warehouse. This partition can be idle for days at a time but will get very busy during the next few weeks. Eventually, at least one-half of a processor will be the desired maximum.

5.3 How it was done

Following the recommendations in Chapter 3, “Setting up LPAR limits to make tuning possible” on page 9, minimum values and maximum values for the processing resources were set as shown in Table 5-2 through Table 5-6 on

Archived

page 34. Because we intend in this example to illustrate both iSeries and System i methodologies, we list both iSeries-centric metrics (including interactive performance percentage) and System i-centric ones (including desired amounts).

5.3.1 For an iSeries system

There are four metrics to specify for iSeries systems: processor power, virtual processors, interactive performance, and memory. For each metric, there are two parameters: a minimum amount and a maximum amount.

Because there is no penalty for setting processor limits as wide as there is hardware capability for every partition, all the partitions were set to a minimum of 0.10 shared processing units and a maximum of 16.00, despite the fact that not all 16 processors were activated. This allows the greatest possible latitude in making later adjustments. This setting also acknowledges the possibility that more processors might be activated using Processor On Demand.

Setting the limits

The limited set for our example is shown in Table 5-2.

Table 5-2 Processor power - iSeries

Partition #	Minimum	Maximum
0 (primary partition)	0.10	16.00
1	0.10	16.00
2	0.10	16.00
3	0.10	16.00
4	0.10	16.00
5	0.10	16.00
6	0.10	16.00
7	0.10	16.00
8	0.10	16.00
Totals	0.90	144.00

Note that the sum of the minimum values (0.90 processing units) is within the capability of the system (16.00 processing units total, 9.00 processing units currently active). This is an obvious requirement for any configuration for processor power, interactive percentage, and memory. This requirement does

Archived

not apply to virtual processors, however. Consider that for a single processor machine with several partitions, at least one virtual processor must be assigned to each partition. The sum of the maximum values far exceeds the capability of the system, but that is reasonable (not all of the maximums can be realized simultaneously, however).

For virtual processors, the minimum in any partition is one (1), and this is the maximum number of tasks that the operating system will attempt to dispatch simultaneously. It makes little sense to set this value below the number of actual processors that will serve the partition at any instant of time. You can set the number of virtual processors to more than the number of processors, but for iSeries hardware, this is usually undesirable from a performance point of view. Because these are not actual values but only limits, and a change to limits requires at least a partition restart (IPL), the limits for all partitions are set as a minimum of 1 and a maximum of 16 as shown in Table 5-3.

Table 5-3 Virtual processors - iSeries

Partition #	Minimum	Maximum
0 (primary partition)	1	16
1	1	16
2	1	16
3	1	16
4	1	16
5	1	16
6	1	16
7	1	16
8	1	16
Totals	9	144

Because this is an iSeries system, limits must be set for interactive performance percentage. Again, there is no penalty for setting the minimum very low and the maximum very high for every partition, so these limits were set at the values shown in Table 5-4 on page 32.

Table 5-4 Interactive processing percentage - iSeries

Partition #	Minimum	Maximum
0 (primary partition)	0	100

Archived

Partition #	Minimum	Maximum
1	0	100
2	0	100
3	0	100
4	0	100
5	0	100
6	0	100
7	0	100
8	0	100
Totals	0	900

Setting limits for memory is different, however. The maximum memory allowed in a partition controls the amount of memory reserved for the hardware page table. It is wasteful to set all the limits to the maximum installed memory in most cases, because it is unlikely that you will ever want some of these partitions to use that much memory. Management, therefore, walked a carefully considered line between flexibility (wide limits with large maximum values) and conservation of resources (reasonable upper limits). Because of the business use of each partition and its relative importance to the management of the company, memory limits were set as shown in Table 5-5 on page 33.

Table 5-5 Memory - iSeries

Partition #	Minimum	Maximum
0 (primary partition)	600 MB	1,000 MB
1	1,000 MB	25,000 MB
2	1,000 MB	14,000 MB
3	1,000 MB	4,000 MB
4	1,000 MB	8,000 MB
5	1,000 MB	4,000 MB
6	1,000 MB	12,000 MB
7	1,000 MB	5,000 MB
8	600 MB	4,000 MB

Archived

Partition #	Minimum	Maximum
Totals	8,200 MB	77,000 MB

Again, the total of the minimum values is realizable with the installed hardware; the maximum is not. Because this example is for an iSeries system, the smallest increment of memory allocation is 1 MB, regardless of the total amount of installed memory.

Setting the actual values

It is also required that you set actual (current) values for all of these metrics for all partitions. The actual values must respect the limits established and, of course, must be realizable with the installed and activated hardware of the system. It is, of course, possible to change these current values to meet business goals from time to time, but each change requires effort and time. The worst change in this regard is to memory. Because there is rarely a good reason for leaving any processing resource unassigned and therefore idle, it is often necessary to remove resource from one or more partitions in order to reassign it to a single partition or other partitions. As we discussed in Chapter 3, “Setting up LPAR limits to make tuning possible” on page 9, removing memory from an active partition takes time and, perhaps, resetting a system value and the reorganization of memory pool allocations. For this reason, carefully planning these values is important.

After due consideration, the management of the company decided on the actual values as shown in Table 5-6 on page 34.

Table 5-6 Actual values - iSeries

Partition #	Processor units	Virtual processors	Interactive performance percent	Memory
0 (primary partition)	0.15	1	1	600 MB
1	5.00	5	31	19,552 MB
2	1.75	2	11	9,000 MB
3	0.25	1	2	2,000 MB
4	0.30	1	2	4,000 MB
5	0.20	1	1	2,000 MB

Archived

Partition #	Processor units	Virtual processors	Interactive performance percent	Memory
6	1.00	1	6	8,000 MB
7	0.25	1	2	3,000 MB
8	0.10	1	1	1,000 MB
Totals	9.00	14	57	49,152 MB

These values use exactly all of the currently active processor units, interactive performance, and memory of the system.

From a business perspective, the most important issue was performance in Partition #1 and Partition #2, which run the two largest business units' workload. With this configuration, a full day graph of processor utilization (the major bottleneck for both workloads) showed that there were periods during the day when these partitions were both rather busy, and there were periods when only one seemed to slow down. Although the total throughput might have benefited from the movement of processing power from one of these partitions to the other partition or, because other partitions were relatively idle during some of the busiest times for these business units, from temporarily reassigning some of their idle resources, this was not done because of the inconvenience and cost. A person needed to recognize the situation and then rearrange the processing power several times over the course of the business day. There is also the possibility of human error in this activity.

The assigned processor levels and percentage utilization for Partition #1 and Partition #2 are shown in Figure 5-1 and Figure 5-2. The vertical bars represent processor utilization; the thick top line represents configured processing units.

Archived

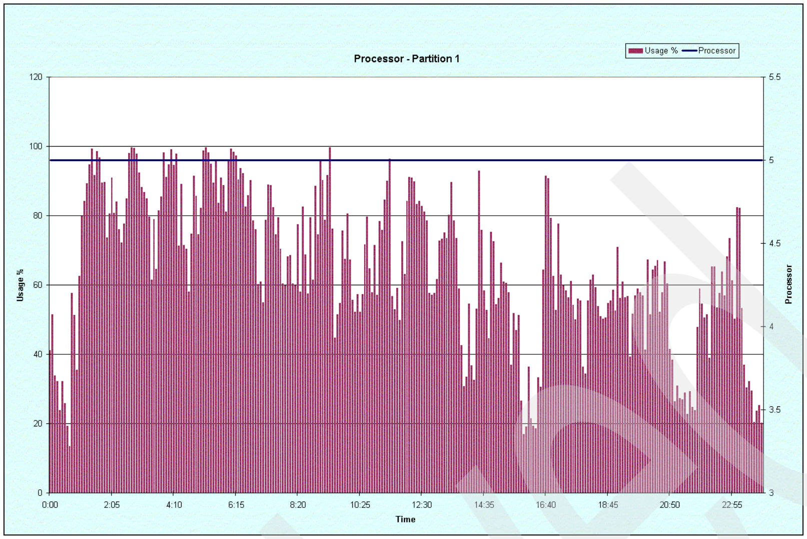


Figure 5-1 Processor utilization and configuration - Partition #1

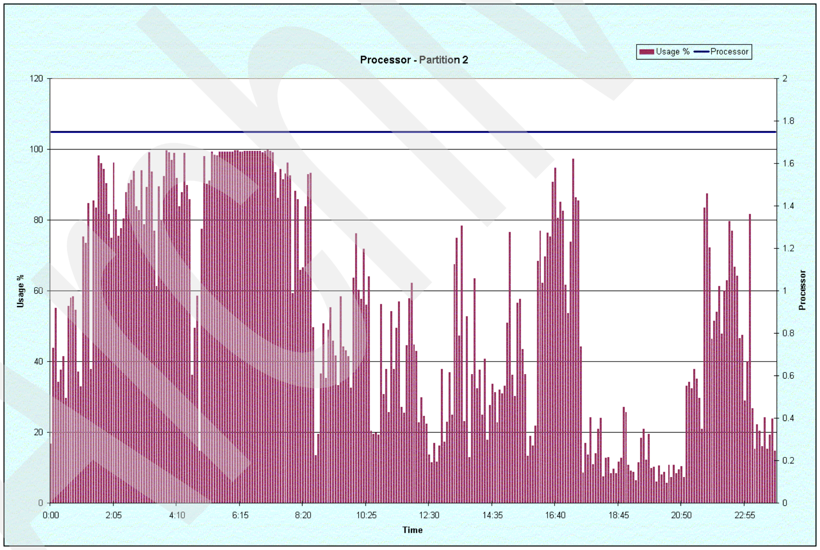


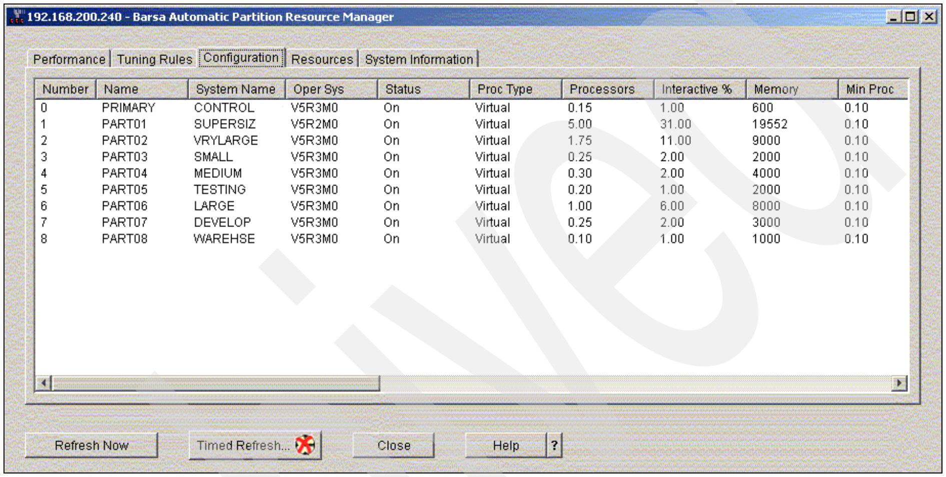
Figure 5-2 Processor utilization and configuration - Partition #2

As you can see in Figure 5-1 and Figure 5-2, Partition #2 (with 1.75 processor units) is starved for processor power for much of the interval from about 5:00 a.m. through about 7:30 a.m. During that same time period, Partition #1 (with 5.00 processor units) is busy but does not use all of its assigned processor power for much of this time. For the particular day shown, Partition #1 can share

Archived

perhaps as much as 1.00 to 1.50 processor units with Partition #2 for this interval and therefore improve the system's total throughput. We revisit this same interval, though not from this day, later in Chapter 6, "Automatic LPAR tuning with APRM" on page 43, with APRM helping productivity by reassigning resources concurrently and appropriately.

You can view the static configuration that we described all at once by using the APRM iSeries Navigator plug-in on the Configuration tab as shown in Figure 5-3.



Number	Name	System Name	Oper Sys	Status	Proc Type	Processors	Interactive %	Memory	Min Proc
0	PRIMARY	CONTROL	V5R3M0	On	Virtual	0.15	1.00	600	0.10
1	PART01	SUPERSIZ	V5R2M0	On	Virtual	5.00	31.00	19552	0.10
2	PART02	VRYLARGE	V5R3M0	On	Virtual	1.75	11.00	9000	0.10
3	PART03	SMALL	V5R3M0	On	Virtual	0.25	2.00	2000	0.10
4	PART04	MEDIUM	V5R3M0	On	Virtual	0.30	2.00	4000	0.10
5	PART05	TESTING	V5R3M0	On	Virtual	0.20	1.00	2000	0.10
6	PART06	LARGE	V5R3M0	On	Virtual	1.00	6.00	8000	0.10
7	PART07	DEVELOP	V5R3M0	On	Virtual	0.25	2.00	3000	0.10
8	PART08	WAREHSE	V5R3M0	On	Virtual	0.10	1.00	1000	0.10

Figure 5-3 iSeries Configuration displayed by APRM iSeries Navigator plug-in

5.3.2 For a System i platform

There are three metrics to specify for this System i platform: processor power, virtual processors, and memory. For each metric, there are three parameters: a minimum amount, a desired amount, and a maximum amount. You specify the desired amount in the partition profile, and this amount is in place of the current or actual setting in System Service Tools (SST), Dedicated Service Tools (DST), or iSeries Navigator. It is possible to create multiple partition profiles for each defined partition to preplan for different start-up scenarios, but we do not show that complication here. We only show one profile for each partition, and we designate that profile as the default profile for simplicity.

Because there is no penalty for setting processor limits as wide as there is hardware capability for every processor, we set all the partitions to a minimum of 0.10 shared processing units and a maximum of 8.00, despite the fact that not all eight processors were activated. This allows the widest possible latitude in making later adjustments. This also acknowledges the possibility that more processors will be activated using Processor On Demand.

Archived

Setting the limits and the desired values

Table 5-7 shows the partition limits and values set for desired processors.

Table 5-7 Processor power - System i

Partition #	Minimum	Desired	Maximum
1	0.10	2.75	8.00
2	0.10	1.00	8.00
3	0.10	0.14	8.00
4	0.10	0.18	8.00
5	0.10	0.10	8.00
6	0.10	0.56	8.00
7	0.10	0.17	8.00
8	0.10	0.10	8.00
Totals	0.80	5.00	64.00

Note that the sum of the minimum values (0.80 processing units) is within the capability of the system (8.00 processing units total, 5.00 processing units currently active). This is an obvious requirement for any configuration for processor power and memory. This requirement does not apply to virtual processors, however. Consider that for a single processor machine with several partitions, there must be at least one virtual processor assigned to each partition. The sum of the maximum values far exceeds the capability of the system, but that is reasonable (not all of the maximums can be realized simultaneously, however). The sum of the desired values has been made to use exactly all of the currently activated processor power, because the intention is to have all the partitions active most of the time. Had this resource been overcommitted, one or more partitions would not have been able to obtain their desired values when started (if all the partitions were activated at once, it would be a race to determine which partition is shortchanged). However, there might have been enough processor power to meet the minimum value and thus allow the partition to be activated.

Because this System i platform supports uncapped processors and management chose to use this capability, you must also specify the *weight* assigned to each partition. The weight is a metric that controls the relative share of any idle processor power that will be assigned to a partition that has already used its own assigned processing power. This assignment is on a 10 millisecond rotation basis; values can be set from zero to 255. A weight of zero means that no fraction of idle processor power will be assigned; this is essentially the same as

Archived

specifying that the partition does not use uncapped processors. Because of the relative importance of the partitions to the company and their relative sizes, management decided to set the weights as shown in Table 5-8.

Table 5-8 Uncapped processor weights - System i

Partition #	Expected processor power assigned	Uncapped weight
1	2.75	150
2	1.00	150
3	0.14	100
4	0.18	100
5	0.10	50
6	0.56	100
7	0.17	100
8	0.10	125
Totals	5.00	875

The total of the uncapped processor weights does not have a practical limit; only the relative weights of those partitions actively vying for otherwise-idle processor cycles make any difference. As with iSeries systems, the minimum in any partition is one (1) for virtual processors. It makes little sense to set this value below the number of actual processors that will serve the partition at any instant of time. You can set the number of virtual processors to more than the number of processors. For System i hardware with the capability for uncapped processors, this can be a desirable thing to do. See the discussion of uncapped processors in 3.3, “System i platforms” on page 16, for the rationale in this regard.

Because this particular System i platform does have the capability to support uncapped processors, and there might eventually be eight activated processors, the maximums were all set to eight virtual processors. Because the current number of activated processors was five, and all the partitions were configured to use processing power left unused by other partitions, the desired value for virtual processors was set at five for all the partitions as shown in Table 5-9 on page 40. This allows any partition to effectively use the entire processing power of the system so long as no other partition has any processor activity.

Archived

Archived

Table 5-9 Virtual processors - System i

Partition #	Minimum	Desired	Maximum
1	1	5	8
2	1	5	8
3	1	5	8
4	1	5	8
5	1	5	8
6	1	5	8
7	1	5	8
8	1	5	8
Totals	8	40	64

Setting limits for memory is different, however. The maximum memory that is allowed in a partition controls the amount of memory reserved for the hardware page table. It is wasteful to set all the limits to the maximum installed memory in most cases because it is unlikely that you will ever want some of these partitions to use that much memory. Management, therefore, walked a carefully considered line between flexibility (wide limits with large maximum values) and conservation of resources (reasonable limits).

Another factor to consider when assigning memory values on System i platforms is the granularity of memory. In order to keep the hardware page table to a manageable size, especially with systems with a terabyte or more of memory, IBM takes the approach of allocating memory only in *regions,* whose minimum size is determined by the total memory installed in the system. You can use a higher value, but we show the minimum values for several total installed memory sizes in Table 5-10 on page 40.

Table 5-10 Memory block sizes compared to configurable memory

Amount of configurable memory	Default memory block size
Less than 4 GB	16 MB
Greater than 4 GB up to 8 GB	32 MB
Greater than 8 GB up to 16 GB	64 MB
Greater than 16 GB up to 32 GB	128 MB
Greater than 32 GB	256 MB

Because this new system had 64 GB of memory, the region size was 256 MB. Because of the business use of each partition and its relative importance to the management of the company, memory limits were set as shown in Table 5-11.

Table 5-11 Memory - System i

Partition #	Minimum	Desired	Maximum
1	2 GB	22 GB	36 GB
2	2 GB	13 GB	24 GB
3	2 GB	3 GB	6 GB
4	2 GB	5 GB	12 GB
5	2 GB	3 GB	6 GB
6	2 GB	10 GB	16 GB
7	2 GB	4 GB	8 GB
8	2 GB	2 GB	4 GB
Totals	16 GB	62 GB	112 GB

Again, the total of the minimum values is realizable with the installed hardware; the maximum is not. The desired values do not use exactly the installed 64 GB of memory; instead, 2 GB were purposefully not included to provide for the memory needed by the SP. If, after the company first brought up the system with all of its partitions activated, this amount proved to be too high or too low, adjustments must be made to one or more of the desired values.

Because this System i platform is a hypothetical configuration, we do not provide any graphs of actual performance.

Archived

Archived

6

Automatic LPAR tuning with APRM

In the previous chapter, we presented and analyzed the partitioning of an iSeries system in a real-life situation. We also presented a hypothetical example of that same scenario with a System i platform in place of the iSeries platform.

This chapter illustrates how those systems might benefit from having an automated product to adjust the logical partitioning (LPAR) configuration according to the varying workloads.

Archived

© Copyright IBM Corp. 2007. All rights reserved.

6.1 General considerations

For each of the LPAR configurations, you must establish a set of Automatic Partition Resource Manager (APRM) tuning rules. These rules differ from the rules provided to the LPAR hypervisor through Dedicated Service Tools (DST), System Service Tools (SST), iSeries Navigator (for iSeries hardware), or the Hardware Management Console (HMC) (for System i hardware) in a few ways:

- These rules must not violate the IBM LPAR limits.
- You can change these rules at any time, and the changes are immediately effective.
- The APRM tuning rules really control the management of the system; the LPAR limits simply set outer boundaries for the APRM rules.
- There is no need nor any capability to specify virtual processors because APRM automatically sets this metric according to the number of processors assigned to the partition. APRM assigns a virtual processor to each full processor or any fraction of a processor so that processing power is not wasted (APRM considers uncapped processors).
- Either simple 5250 application Command Language (CL) commands, a 5250 application interactive application, or an iSeries Navigator plug-in graphical user interface (GUI) can retrieve, view, or change these rules.
- Because APRM provides 5250 application CL commands that can change the APRM tuning rules, any job scheduler including the standard Add Job Schedule Entry (ADDJOBSCDE), Change Job Schedule Entry (CHGJOBSCDE), Remove Job Schedule Entry (RMVJOBSCDE), and Work with Job Schedule Entries (WRKJOBSCDE) commands can change these rules according to a time schedule.

For the real-life iSeries system illustrated in Chapter 5, “A real-life practical example of LPAR configuration” on page 27, we present real results for the time period of greatest resource contention.

6.2 iSeries system tuning rules

For an iSeries system, you must specify three metrics: processor power, interactive performance, and memory. We discuss tuning rule considerations for processing power and interactive performance separately by system type. Memory limits, although not restricted by other resources, can affect tuning performance because of the amount of time that it can require to remove memory from an active partition. Before memory can be removed from an OS/400 or i5/OS partition, it must be moved through the machine pool, its

Archived

content saved to disk, and then moved to the base pool from which the LPAR hypervisor can take it.

Even with caching disk controllers, it can still take several minutes to empty a significant block (gigabytes) of memory. While memory is removed from a partition, you cannot make any other adjustments at all. During this time, the page fault rates in the partition that is losing the memory can be greater than the production workload otherwise causes. Also, this action usually requires a reallocation of the remaining memory among the shared pools to maintain reasonable performance in the partition. Moreover, when the memory is then added to another partition, it is given to the base pool there. Perhaps another shared pool in that partition can use some of the added memory for efficiency. For these reasons, APRM requires that IBM memory reallocation software is enabled (that is, system value QPFRADJ is set to 2 or 3), and APRM rearranges the shared pools in the partition that is about to lose memory in advance of taking it in order to minimize the impact of the memory loss.

As with the system LPAR limits, set the minimum value for interactive performance and the minimum value for processor power to be compatible. That is, you might not assign so much interactive power that the partition’s assigned processing power cannot meet that minimum. For many systems, this is not an issue because there is insufficient total interactive power to overburden even a minimal processor allocation, but for a system with significant or unrestricted interactive performance, this is a requirement. Because the machine that we configure here is an Enterprise-class machine (no artificial interactive limit), the minimum interactive performance is important.

6.2.1 Setting the limits

For processor power, management decided on the following values shown in Table 6-1 on page 46.

Archived

Table 6-1 APRM limits for processor power - iSeries

Partition #	Minimum	Maximum
0 (primary partition)	0.10	9.00
1	2.50	5.00
2	1.00	3.00
3	0.10	1.50
4	0.10	1.50
5	0.10	1.50
6	0.35	2.75
7	0.10	3.00
8	0.10	0.50
Totals	4.45	27.75

Note that management chose these numbers to limit the full ranges of possible values to those that management wanted to actually use, at least at first. Because you can easily change the APRM tuning rules without disruption of processing, a bad guess here is easy to correct. As always, the sum of the minimum values must be realizable with the installed hardware, and it is. The sum of the maximum values must be higher than the total installed hardware to provide for room to make adjustments. Of course, these limits must always be within the system LPAR limits.

We calculated the limits for interactive performance so that the minimum interactive percentage is compatible with the minimum processor power, and the maximum interactive percentage is compatible with the maximum processor power. The result is shown in Table 6-2 on page 47.

Archived

Table 6-2 APRM limits for interactive processing percentage - iSeries

Partition #	**Minimum**	**Maximum**
0 (primary partition)	1	57
1	16	31
2	6	19
3	1	9
4	1	9
5	1	9
6	2	17
7	1	19
8	1	3
Totals	30	173

We initially set up the tuning rules for memory as shown in Table 6-3 on page 48, Figure 6-1 on page 49, and Figure 6-2 on page 50. Because, as we discussed in Chapter 3, “Setting up LPAR limits to make tuning possible” on page 9 and Chapter 4, “Setting up APRM limits to make tuning efficient” on page 19, memory moves can take significant time, it proved impractical to allow memory changes throughout the business day. The result was that, in this installation, processor power moves were delayed too long by waiting for memory moves to complete. In response to this, management of the company decided to move memory only at specifically defined times of the day in advance of expected workload shifts. These changes were also performed by APRM, but the means employed was to change the tuning rules (just for memory values) and, at each change, to make the minimum memory values and maximum memory values the same. We implemented these changes using Change APRM Tuner (CHGAPRMTNR) commands first through the standard job scheduler entries and then, later, by a more flexible CL program that had logic to respond to weekends, holidays, month-end processing needs, and the time of day. These are not illustrated here.

Archived

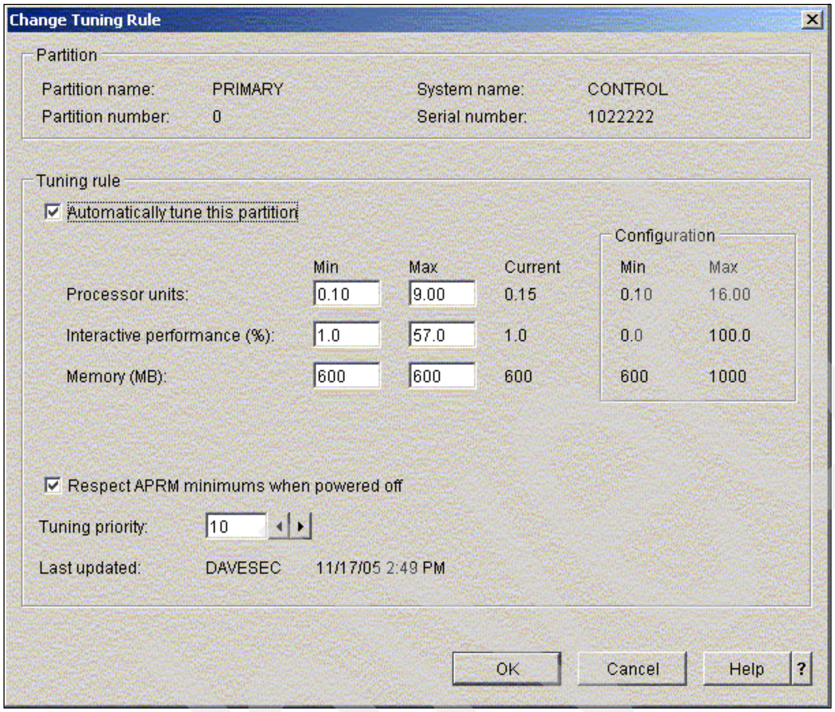
With that said, we show the original memory tuning rules in Table 6-3.

Table 6-3 APRM limits for memory - iSeries

Partition #	Minimum	Maximum
0 (primary partition)	600 MB	600 MB
1	14,000 MB	20,000 MB
2	4,000 MB	12,000 MB
3	1,000 MB	3,000 MB
4	1,500 MB	6,500 MB
5	1,000 MB	3,000 MB
6	3,500 MB	10,000 MB
7	1,500 MB	4,000 MB
8	1,000 MB	3,000 MB
Totals	28,100 MB	62,100 MB

You can enter these values interactively through the 5250 application Work with APRM (WRKAPRM) command (not illustrated) or through the APRM iSeries Navigator plug-in as shown in Figure 6-1 on page 49.

Archived



Change Tuning Rule

Partition

Partition name: PRIMARY
System name: CONTROL
Partition number: 0
Serial number: 1022222

Tuning rule

☑ Automatically tune this partition

	Min	Max	Current	Configuration Min	Configuration Max
Processor units:	0.10	9.00	0.15	0.10	16.00
Interactive performance (%):	1.0	57.0	1.0	0.0	100.0
Memory (MB):	600	600	600	600	1000

☑ Respect APRM minimums when powered off

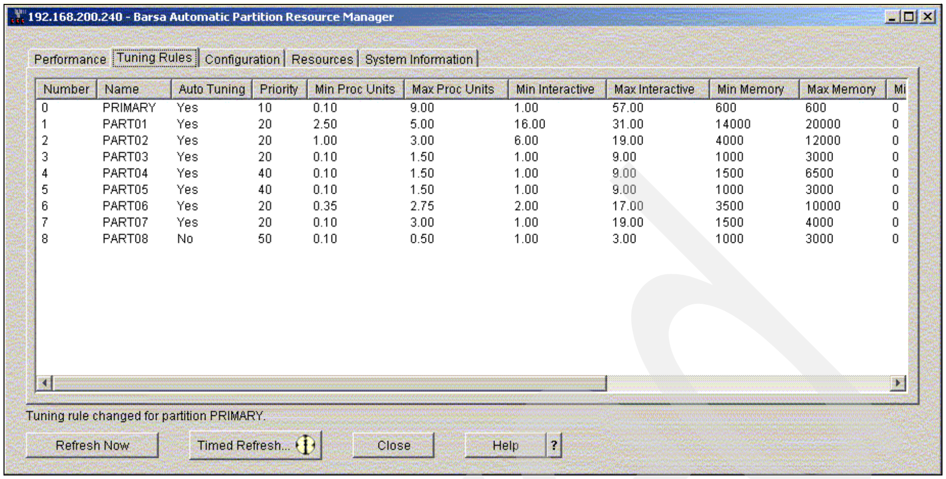
Tuning priority: 10

Last updated: DAVESEC 11/17/05 2:49 PM

Figure 6-1 Change tuning rules using GUI plug-in

You can view the tuning rules through the plug-in tab for that purpose as shown in Figure 6-2 on page 50. You can resize the window, resize the columns, and reorder the columns as you want. There are more columns than we show in Figure 6-2 on page 50, but those shown include the rules for all nine partitions and all three controlled resources in one display.

Archived



192.168.200.240 - Barsa Automatic Partition Resource Manager

Number	Name	Auto Tuning	Priority	Min Proc Units	Max Proc Units	Min Interactive	Max Interactive	Min Memory	Max Memory	Mi
0	PRIMARY	Yes	10	0.10	9.00	1.00	57.00	600	600	0
1	PART01	Yes	20	2.50	5.00	16.00	31.00	14000	20000	0
2	PART02	Yes	20	1.00	3.00	6.00	19.00	4000	12000	0
3	PART03	Yes	20	0.10	1.50	1.00	9.00	1000	3000	0
4	PART04	Yes	40	0.10	1.50	1.00	9.00	1500	6500	0
5	PART05	Yes	40	0.10	1.50	1.00	9.00	1000	3000	0
6	PART06	Yes	20	0.35	2.75	2.00	17.00	3500	10000	0
7	PART07	Yes	20	0.10	3.00	1.00	19.00	1500	4000	0
8	PART08	No	50	0.10	0.50	1.00	3.00	1000	3000	0

Tuning rule changed for partition PRIMARY.

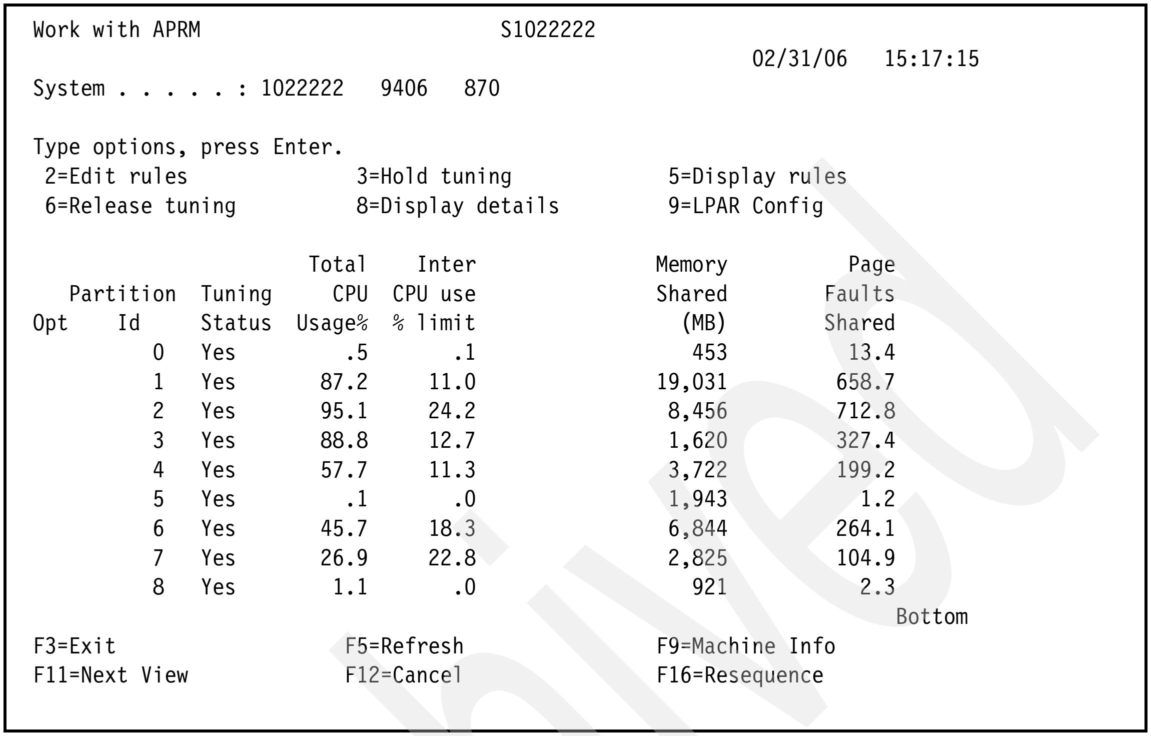
Figure 6-2 Tuning rules tabular display using GUI plug-in

6.2.2 How well did it work and how do you know

There are several ways to view how well or how poorly your tuning rules and APRM perform at keeping your workload moving along. Two of these ways, the 5250 application Work with APRM (WRKAPRM) command display and the Performance tab of the iSeries Navigator APRM plug-in, display a current picture of processor usage, page fault rates (memory usage), and interactive usage compared to the amount allowed by your hardware. There is also a Convert APRM History (CVTAPRMHST) command that extracts historical data backward in time from journal entries over the configuration and performance data files and then writes the analyzed data to files in the Integrated File System (IFS). These files are in comma-separated value (CSV) format. You can analyze these files by several standard data presentation and reporting tools. APRM also provides a Microsoft® Excel® add-in that allows you to easily display this data graphically.

The Work with APRM (WRKAPRM) command provides a 5250 application interface to display partition performance, edit and view APRM tuning rules, hold and release APRM tuning, view the LPAR configuration, and view the machine's configuration. You can "window" the partition performance display by repeatedly pressing F11, and you can change the order of the columns. At one moment in time, the display appeared as shown in Figure 6-3 on page 51.

Archived

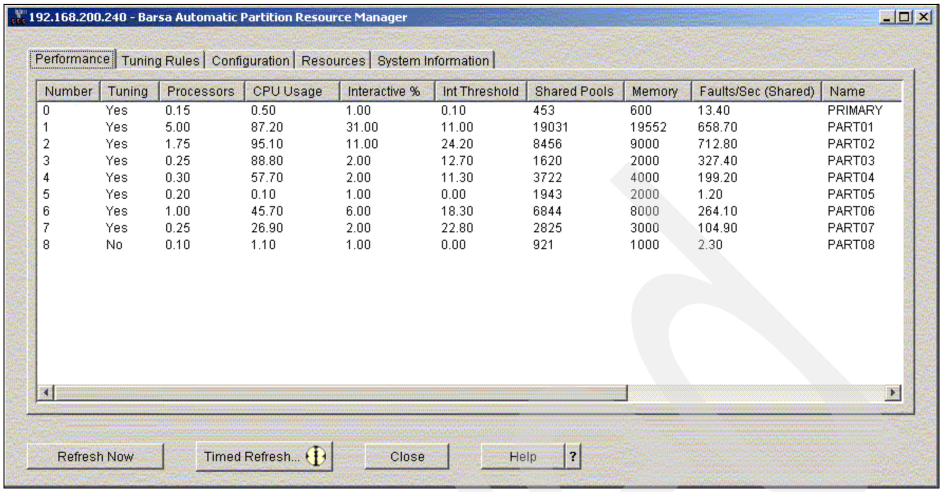


```
Work with APRM                       S1022222
                                                        02/31/06   15:17:15
System . . . . . : 1022222   9406   870

Type options, press Enter.
 2=Edit rules             3=Hold tuning            5=Display rules
 6=Release tuning         8=Display details        9=LPAR Config

                       Total    Inter              Memory          Page
   Partition  Tuning     CPU  CPU use              Shared        Faults
Opt    Id     Status  Usage%  % limit                (MB)        Shared
         0    Yes         .5       .1                 453          13.4
         1    Yes       87.2     11.0              19,031         658.7
         2    Yes       95.1     24.2               8,456         712.8
         3    Yes       88.8     12.7               1,620         327.4
         4    Yes       57.7     11.3               3,722         199.2
         5    Yes         .1       .0               1,943           1.2
         6    Yes       45.7     18.3               6,844         264.1
         7    Yes       26.9     22.8               2,825         104.9
         8    Yes        1.1       .0                 921           2.3
                                                                      Bottom
F3=Exit                  F5=Refresh                F9=Machine Info
F11=Next View            F12=Cancel                F16=Resequence
```

Figure 6-3 Work with APRM 5250 application display

You can view this same data on the Performance tab of the iSeries Navigator plug-in, as shown in Figure 6-4 on page 52. Here, you can resize the columns and reorganize them to suit your needs. You can also sort the data by the instant value in any column by clicking on the column heading.

Archived



Number	Tuning	Processors	CPU Usage	Interactive %	Int Threshold	Shared Pools	Memory	Faults/Sec (Shared)	Name
0	Yes	0.15	0.50	1.00	0.10	453	600	13.40	PRIMARY
1	Yes	5.00	87.20	31.00	11.00	19031	19552	658.70	PART01
2	Yes	1.75	95.10	11.00	24.20	8456	9000	712.80	PART02
3	Yes	0.25	88.80	2.00	12.70	1620	2000	327.40	PART03
4	Yes	0.30	57.70	2.00	11.30	3722	4000	199.20	PART04
5	Yes	0.20	0.10	1.00	0.00	1943	2000	1.20	PART05
6	Yes	1.00	45.70	6.00	18.30	6844	8000	264.10	PART06
7	Yes	0.25	26.90	2.00	22.80	2825	3000	104.90	PART07
8	No	0.10	1.10	1.00	0.00	921	1000	2.30	PART08

Figure 6-4 Performance tab of the APRM iSeries Navigator plug-in

Historical data provides another insight into performance tuning. We charted the usage history for two partitions over an eight-hour period, and we discussed this history in Chapter 5, “A real-life practical example of LPAR configuration” on page 27. This data was taken from a day at a later time than the graphs presented in Chapter 5, “A real-life practical example of LPAR configuration” on page 27; this data illustrates usage after APRM made resource adjustments. Because there were several other busy partitions that wanted some resource during this same time interval, the sum of the processor-assigned values for these two partitions is not a constant. You can see, however, that Partition #2 experienced nearly 100% processor utilization only from about 6:10 a.m. through 7:40 a.m., probably because it had its maximum complement of 3.00 processing units during most of that interval. Partition #1 was reduced from its maximum of 5.00 processing units to near its minimum of 2.50 processing units, which provided enough unused resource to Partition #2, yet Partition #1 was well below 100% usage for all but one five-minute interval.

In Figure 6-5 on page 53 and Figure 6-6 on page 53, the vertical bars represent processor utilization, and the line represents processing units configured.

Archived

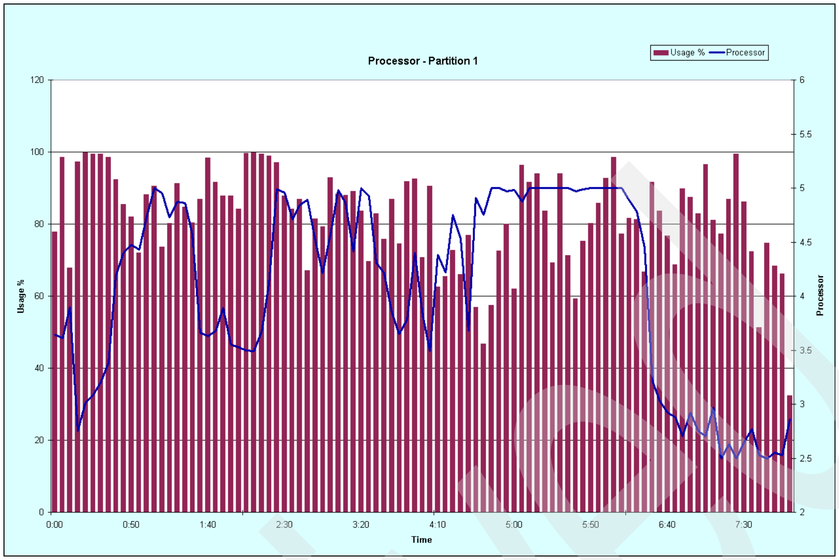


Figure 6-5 Processor utilization and configuration - Partition #1

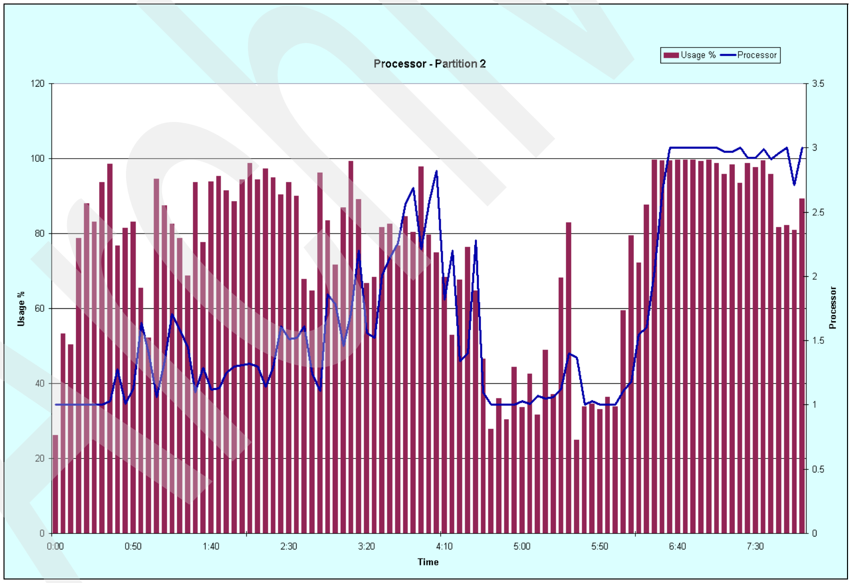


Figure 6-6 Processor utilization and configuration - Partition #2

The use of the APRM automatic tuning algorithm seems to have shortened the interval for Partition #2 to complete its heavy morning workload without any apparent harm to Partition #1. Management later considered changing Partition

Archived

#2's maximum processor limit to more than 3.00 processing units as a result of this analysis.

6.3 System i platform tuning rules

For this System i platform, you must specify only two metrics: processor power and memory; you cannot control interactive performance on this hardware. It is an Enterprise class machine, however, so there is nothing, except for the total available processor capability, to limit interactive use. If this was one of the (mostly small) System i platforms with limited but not zero interactive performance, you must assign the interactive performance percentage in units of 0.01% (that is, 0.00 through 100.00%).

6.3.1 Setting the limits

This system, as defined in Chapter 5, "A real-life practical example of LPAR configuration" on page 27, had eight processors installed, but only five were activated. Remember that each processor here provides more computing power than each of the 16 processors in the iSeries system that it replaces. It also has 64 GB of installed memory, but some must be reserved for the Service Processor (SP), which is system firmware.

For processor power, management decided on the following values shown in Table 6-4.

Table 6-4 APRM limits for processor power - System i

Partition #	Minimum	Maximum
1	1.00	3.00
2	0.40	2.00
3	0.10	1.00
4	0.10	1.00
5	0.10	1.00
6	0.20	2.00
7	0.10	2.00
8	0.10	0.50
Totals	2.10	12.50

Archived

Note that we chose these numbers to limit the full ranges of possible values to those that management wanted to actually use, at least at first. Because you can easily change the APRM tuning rules without disruption of processing, a bad guess here is easy to correct. As always, the sum of the minimum values must be realizable with the installed hardware, and it is. The sum of the maximum values must be higher than the total installed hardware to provide for room to make adjustments. Of course, these limits must always be within the IBM LPAR limits.

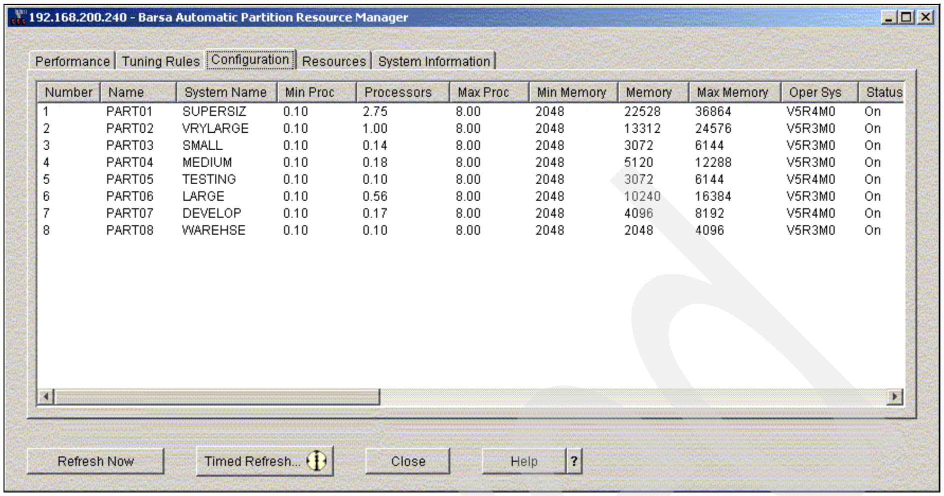
As with the iSeries system that we discussed earlier, you must establish APRM limits for memory. We show these limits in Table 6-5. Whether you adjust memory frequently during the day is a management decision; the rules shown assume that memory will be tuned. Because you can change APRM tuning rules easily and immediately, you can make this decision by using a trial-and-error process. Simply change the minimum values and maximum values to be the same, and APRM has nothing to adjust. Because this is a System i platform, every memory value must be an integer multiple of the memory "region size," which is 256 MB for this system.

Table 6-5 APRM limits for memory - System i

Partition #	Minimum	Maximum
1	14,336 MB	23,040 MB
2	4,096 MB	15,360 MB
3	2,048 MB	3,072 MB
4	2,048 MB	7,168 MB
5	2,048 MB	3,840 MB
6	4,096 MB	12,032 MB
7	2,048 MB	6,144 MB
8	2,048 MB	4,096 MB
Totals	32,768 MB	74,752 MB

You can view the LPAR configuration with its limits and current values through the APRM iSeries Navigator plug-in on the Configuration tab as shown in Figure 6-7 on page 56. You can use the tuning rules tab (not illustrated) to show the APRM rules for all of these partitions in one display.

Archived



Number	Name	System Name	Min Proc	Processors	Max Proc	Min Memory	Memory	Max Memory	Oper Sys	Status
1	PART01	SUPERSIZ	0.10	2.75	8.00	2048	22528	36864	V5R4M0	On
2	PART02	VRYLARGE	0.10	1.00	8.00	2048	13312	24576	V5R3M0	On
3	PART03	SMALL	0.10	0.14	8.00	2048	3072	6144	V5R3M0	On
4	PART04	MEDIUM	0.10	0.18	8.00	2048	5120	12288	V5R3M0	On
5	PART05	TESTING	0.10	0.10	8.00	2048	3072	6144	V5R4M0	On
6	PART06	LARGE	0.10	0.56	8.00	2048	10240	16384	V5R3M0	On
7	PART07	DEVELOP	0.10	0.17	8.00	2048	4096	8192	V5R4M0	On
8	PART08	WAREHSE	0.10	0.10	8.00	2048	2048	4096	V5R3M0	On

Figure 6-7 System i Configuration displayed by APRM iSeries Navigator plug-in

Because this System i platform is a hypothetical configuration, no graphs of actual performance are available.

Archived

7

Sharing I/O hardware

Switching input/output processor (IOP) resources among partitions provides an effective method of sharing hardware resources, such as tape or CD drives. With Automatic Partition Resource Manager (APRM), you can move IOP resources using either an i5/OS command or programmatically.

Archived

© Copyright IBM Corp. 2007. All rights reserved.

7.1 General considerations

One or more partitions can share input/output (I/O) hardware, but there are rules: both rules that are imposed by the operating system and System Licensed Internal Code (SLIC) and rules that are reasonable and just good operating procedure. For example, it is impossible to have two or more partitions controlling a single card slot at the same time; the hardware cannot support this situation, nor does it make any sense to configure this situation. Because of the way that iSeries and System i machines use mass storage, also called disk or DASD (Direct Access Storage Device), it is seldom desirable to switch ownership and usage of a card slot that controls DASD between partitions. However, this is not true in the case of a switchable Independent Auxiliary Storage Pool (IASP). Also, typical best practices mandate that a partition is allowed to finish using a piece of I/O hardware and shut down any related activity before taking that hardware away from the partition for use by another partition.

7.2 iSeries systems

In order for a partition to use I/O hardware, that partition must own or share the I/O bus that controls the slots that hold the circuit cards that connect to the I/O hardware. The logical partitioning (LPAR) hypervisor controls and uses this bus. You can access the LPAR hypervisor through the primary partition of the iSeries system either through the iSeries Navigator or the 5250 application service tools, System Service Tools (SST) or Dedicated Service Tools (DST). A programming Application Program Interface (API) is also available in the primary partition to control LPAR. This API, QYHCHCOP, has a number of security restrictions, communications requirements, and its own grammar and protocol for its interaction with user-written programs. You can change the ownership of an I/O bus and which partitions share an I/O bus while any or all partitions are running, and this change is immediately effective.

When one or all of the partitions are powered off (and thus the system, if the primary partition is involved) and then restarted, no resource changes between or among partitions can occur. An operator, through any of these tools, can make manual changes as desired (as long as the primary partition is functioning) for any other partition whether or not that partition is started. It is possible to render a partition unable to start for lack of resource in this process.

Partitions that use OS/400 traditionally require that input/output adapters (IOAs) are controlled by IOPs and that the I/O slots in each tower have designated groups of IOP-capable slots and the specific IOA-capable slots that they control. Partitions using Linux can use all of these slots for IOAs (Linux neither uses nor requires IOPs). There are newer IOAs (for example, the 5706 and 5707 Ethernet

Archived

IOAs) that do not require an IOP for any operating system but will use one if the IOAs are placed in a slot controlled by an IOP.

7.3 System i platforms

Any partition can be given control of any I/O slot that is not currently required by a different partition. There is, in each partition profile, a list of I/O slots that are required and desired; the lack of a slot on this list does not prevent the partition from using the slot if the slot is assigned to the partition.

The Hardware Management Console (HMC) controls the current ownership of an I/O slot, in addition to several other System i functions. First, the partition profile that is used to start each currently active partition can require that certain I/O slots are reserved for that partition. If no other partition requires a specific slot, the HMC can switch ownership of that slot to any active partition.

The HMC is accessible through:

- Its own GUI console
- A browser-based interface, a “fat client” called Web-Based System Manager (often called WebSM) that provides most of the feel and function of the GUI console
- A UNIX-style command-line interface on the HMC console keyboard or by using a Secure Shell (SSH) connection from another computer that can connect to the HMC through a network.

You can change the requirement for an I/O slot in the partition profile that was used to start a partition at any time while any or all partitions are running, and this change is immediately effective.

When a partition is powered down, you cannot directly take any resource away from it through an HMC command. You can, however, simply use some or all of that partition’s desired or needed resources by assigning them to other partitions that are activated. For memory and processor resources, if multiple partitions are inactive, there is little control over which partition will be depleted or to what extent. You can inadvertently make a partition unable to start for lack of resource in this process.

When several partitions list a particular I/O slot as desired and that slot is not assigned to any active partition, you can change ownership of this slot simply by starting one or more of the partitions that desire the slot. If several partitions start at the same time, the first partition that requests the slot gets ownership for this slot. In this manner, a slot cannot be owned by the same partition that had the slot before you power down several partitions or the entire system.

Archived

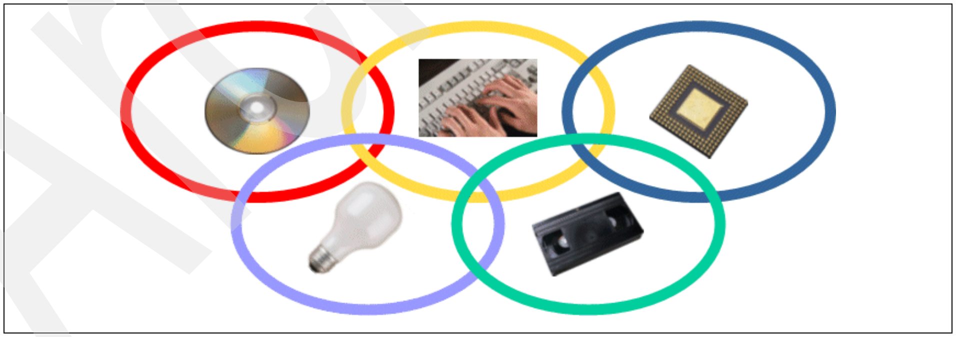
Partitions using i5/OS require many IOAs to be controlled by an IOP and that the I/O slots in each tower have designated groups of IOP-capable slots and the specific IOA-capable slots that they control. Partitions using Linux or AIX can use all of these slots for IOAs (Linux and AIX neither use nor require IOPs).

Archived

8

I want that tape drive right now

In Chapter 7, “Sharing I/O hardware” on page 57, we introduced the iSeries and System i rules and methods for time-sharing input/output (I/O) hardware. This chapter discusses the tools within APRM to make switching hardware easy and safe.



Archived

© Copyright IBM Corp. 2007. All rights reserved.

8.1 General discussion

Automatic Partition Resource Manager (APRM) provides two facilities for switching I/O hardware between partitions. These facilities are termed *manual* and *automatic*, although these terms are not sufficiently descriptive. Both of these facilities operate within rules that you set up to specify which hardware you want to make switchable and, of course, within IBM rules as implemented by the logical partitioning (LPAR) controller. These rules depend upon whether you use iSeries or System i hardware because the LPAR control for iSeries hardware is built into OS/400 and runs in the primary partition. The LPAR control for iSeries hardware uses processor cycles from the primary partition and relies on the primary partition being active. For System i hardware, this is a separate Service Processor (SP) in the System i unit that you must control from the separate Hardware Management Console (HMC), which is an additional computer with its own display and keyboard.

The manual switching facility requires you to assign an APRM name to the controlling input/output processor (IOP) for the input/output adapter (IOA) for the device you want to switch. For System i hardware with IOP-less IOAs, the IOA is the device that you name. In order to make it easy for you to determine the location (slot) of the IOP that controls a device, APRM provides the Display APRM Device (DSPAPRMDEV) command, which gives you the bus, frame number, and card position (slot) of the device's controlling IOP. You must ensure that the partition that owns the IOP and IOA hardware is not using any of the devices, controllers, or lines that rely on the hardware. With the Move APRM Resource (MOVAPRMRSC) command, you can then move the controlling IOP or IOA away from the owning partition to either "available," which means that it is not assigned to any partition, or to another partition. After the partition that receives the hardware has recognized and initialized the hardware, you can then vary *ON any devices that then exist and are usable in the receiving partition. This facility can, from an OS/400 partition, switch hardware to and from Linux, AIX, OS/400 partitions, and the "available" status.

After an APRM name has been assigned, you can use the MOVAPRMRSC command. The command parameters are:

```
MOVAPRMRSC   RSC( IOPname )   TOPTN( partition# / *AVAIL )
```

This command attempts to move *IOPname* to either *partition#* or to *AVAIL (not assigned to any partition).

The "automatic" switching facility requires less user intervention and is suitable for use in an automated backup program. The command for this is Acquire APRM Device (ACQAPRMDEV), in which you specify the name of the device that you want to use in the partition where you execute the command.

Archived

It is the responsibility of APRM to determine which partition (if any) owns the needed I/O hardware, to attempt to get that partition to release it (vary *OFF any devices, controllers, or lines that use the I/O hardware), and to switch the ownership of the I/O hardware to the requesting partition. Then, the Acquire APRM Device (ACQAPRMDEV) support code, optionally, waits for the operating system to recognize and initialize the hardware, and finally varies *ON the requested device. This facility can switch hardware only to and from OS/400 partitions and “available.” The command parameters for the ACQAPRMDEV command are:

```
ACQAPRMDEVDEV( devicename )
        VRYDEVON( *YES | *NO )
        TIMELIMIT( seconds | *NOWAIT )
        RETRYLMT( *NONE | #_of_retries  time_interval )
```

The parameter descriptions are:

- *devicename* specifies the name of the desired device. The device description must exist in the requesting partition.
- *seconds* specifies the desired number of seconds to complete the move and vary on the device.
- *#_of_retries* specifies how many retry attempts to make to acquire the IOP.
- *time_interval* specifies the number of minutes to wait between retry attempts. This command is ideal for a nightly backup job that needs to get the tape drive, vary it *ON, allocate it (so that no other partition can vary it *OFF and take it away until the backup completes), perform the needed backup commands, vary *OFF, and then deallocate the tape drive so that another partition can use it. There is a parameter of the device description for tape devices described as “Unload device at vary off” (UNLOAD), which you can set to leave the tape in a “ready” state if that is what you want. It is still important to consider tape positioning parameters, such as the ENDOPT parameter (*LEAVE or *UNLOAD or *REWIND) and SEQNBR (*number* or *END), but it is possible to automatically time-share an expensive, fast tape device among multiple partitions with the ACQAPRMDEV command without the need for an operator to babysit the process.

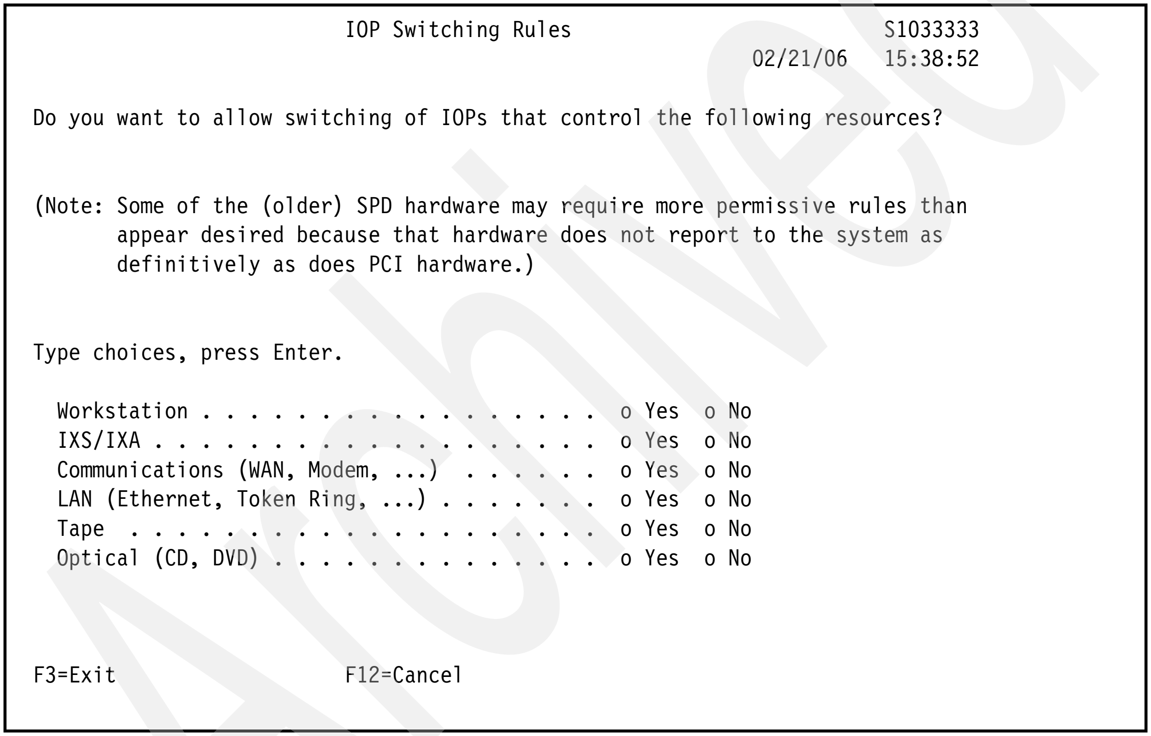
8.2 iSeries systems

The ability to switch an IOP from one partition to another partition presents both advantages and disadvantages to the partition environment. The most obvious advantage is that you can share hardware, such as CD drives or tape drives, among partitions, therefore eliminating the need to purchase hardware for each partition. In contrast, you can inadvertently move hardware resources that can

Archived

threaten the integrity of the system. For example, you do not want to switch an IOP that controlled data and application storage.

To protect your hardware resource environment among partitions, APRM allows you to define what types of hardware are acceptable to move from one partition to another partition. You can designate which kinds of hardware you allow to be switched between partitions and which kinds you do not. You do this with an option (# 5) of the Configure APRM (CFGAPRM) command. Because, on iSeries hardware, it is possible for APRM to determine the type of devices that are connected to an IOA, APRM can enforce these rules. Figure 8-1 shows these switching rule choices.

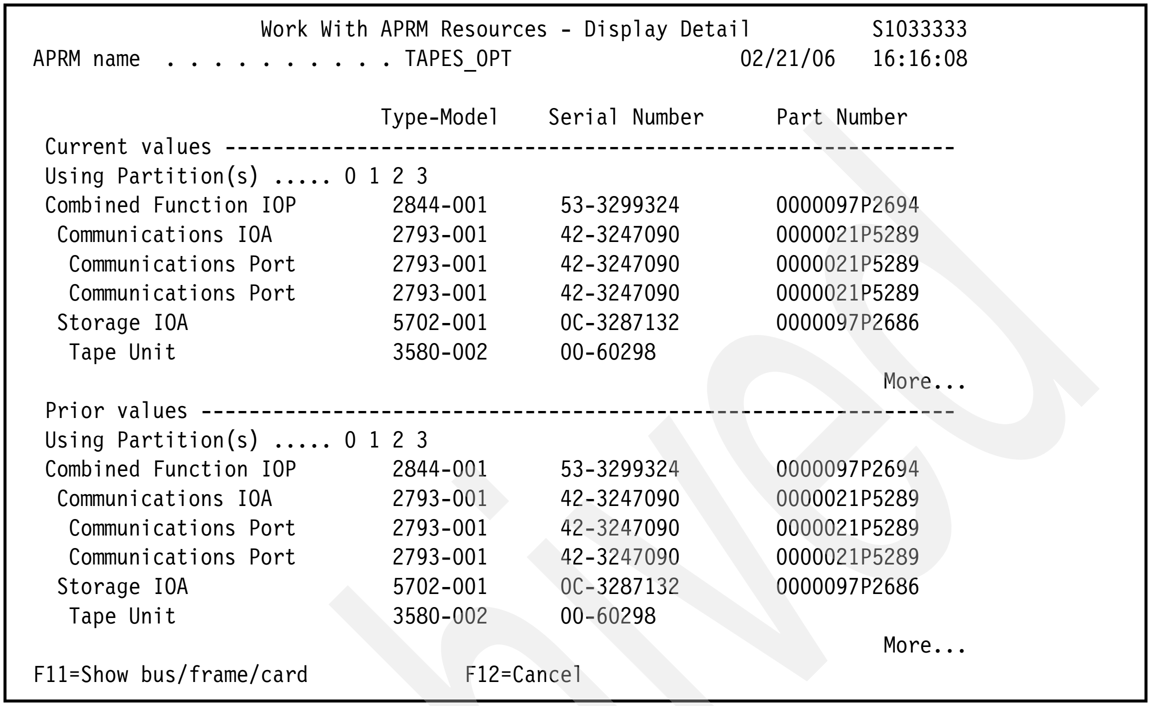


```
                         IOP Switching Rules                          S1033333
                                                             02/21/06   15:38:52

Do you want to allow switching of IOPs that control the following resources?


(Note: Some of the (older) SPD hardware may require more permissive rules than
       appear desired because that hardware does not report to the system as
       definitively as does PCI hardware.)


Type choices, press Enter.

  Workstation . . . . . . . . . . . . . . . . .  o Yes  o No
  IXS/IXA . . . . . . . . . . . . . . . . . . .  o Yes  o No
  Communications (WAN, Modem, ...)  . . . . . .  o Yes  o No
  LAN (Ethernet, Token Ring, ...) . . . . . . .  o Yes  o No
  Tape  . . . . . . . . . . . . . . . . . . . .  o Yes  o No
  Optical (CD, DVD) . . . . . . . . . . . . . .  o Yes  o No



F3=Exit                  F12=Cancel
```

Figure 8-1 Configure APRM 5250 application display

For a partition to be able to use any hardware, the partition must be able to use the I/O bus that connects to the IOPs and IOAs controlling that hardware. The bus must be designated as either “Own Shared” or “Use Shared” by the partition. When you view the list of switchable hardware by using the Work with APRM Resources (WRKAPRMRSC) command and then display details of a specific IOP, you can see a list of which partitions are eligible to use the hardware that the IOP controls. This is illustrated in Figure 8-2 on page 65.

Archived



```
                   Work With APRM Resources - Display Detail          S1033333
APRM name  . . . . . . . . . . TAPES_OPT                   02/21/06   16:16:08

                            Type-Model    Serial Number      Part Number
 Current values -------------------------------------------------------------
 Using Partition(s) ..... 0 1 2 3
 Combined Function IOP       2844-001      53-3299324        0000097P2694
  Communications IOA         2793-001      42-3247090        0000021P5289
   Communications Port       2793-001      42-3247090        0000021P5289
   Communications Port       2793-001      42-3247090        0000021P5289
  Storage IOA                5702-001      0C-3287132        0000097P2686
   Tape Unit                 3580-002      00-60298
                                                                      More...
 Prior values ---------------------------------------------------------------
 Using Partition(s) ..... 0 1 2 3
 Combined Function IOP       2844-001      53-3299324        0000097P2694
  Communications IOA         2793-001      42-3247090        0000021P5289
   Communications Port       2793-001      42-3247090        0000021P5289
   Communications Port       2793-001      42-3247090        0000021P5289
  Storage IOA                5702-001      0C-3287132        0000097P2686
   Tape Unit                 3580-002      00-60298
                                                                      More...
F11=Show bus/frame/card              F12=Cancel
```

Figure 8-2 Work With APRM Resources - Display Detail for iSeries

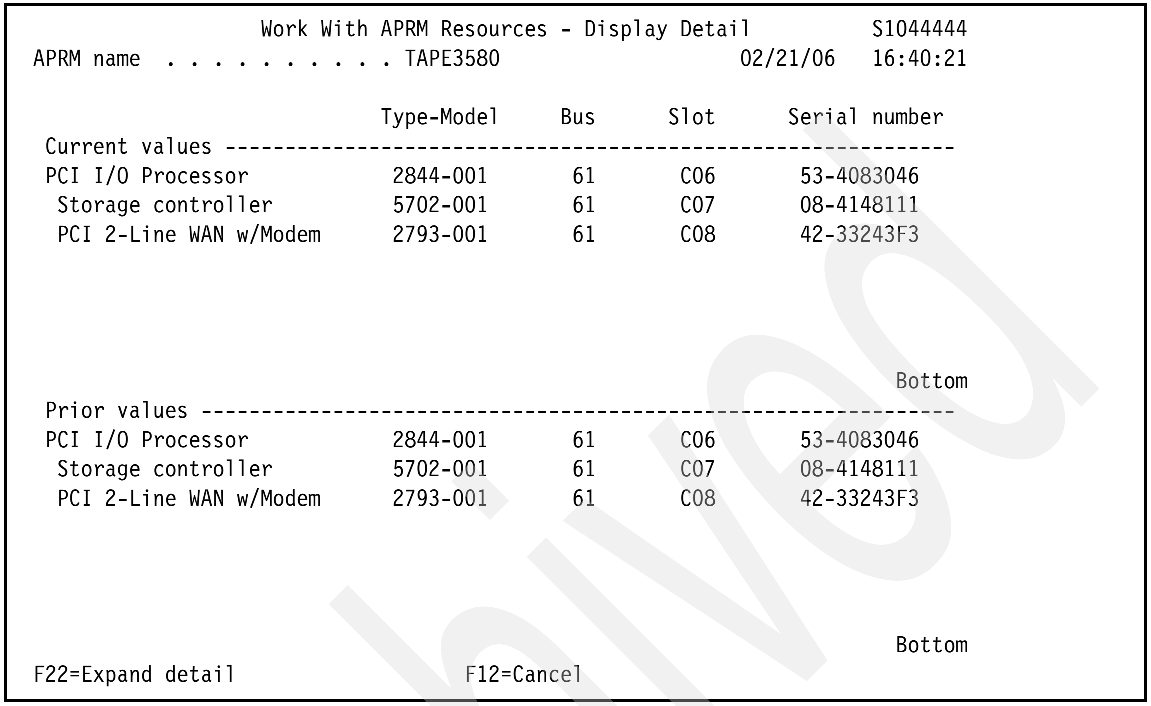
8.3 System i platforms

Because the LPAR controller, the SP, and the HMC do not supply much information below the IOA level, APRM cannot implement rules about which kinds of hardware you want to make switchable. Instead, you assign an APRM name (as you do for the manual switching facility) as the indication to APRM that a specific IOP or IOP-less IOA can be switched.

Any partition can use any slot on any bus so long as the slot is not currently required for another partition. APRM does not take an IOP from a partition that currently uses it (that is, has any devices, controllers, or lines varied *ON that use the IOP or any of its IOAs).

The WRKAPRMRSC detail display (shown as Figure 8-1 on page 64 for iSeries hardware) changes a little with System i hardware. The list of possible partitions is absent (because it is simply ALL of the partitions), and the column headings change as shown in Figure 8-3 on page 66.

Archived

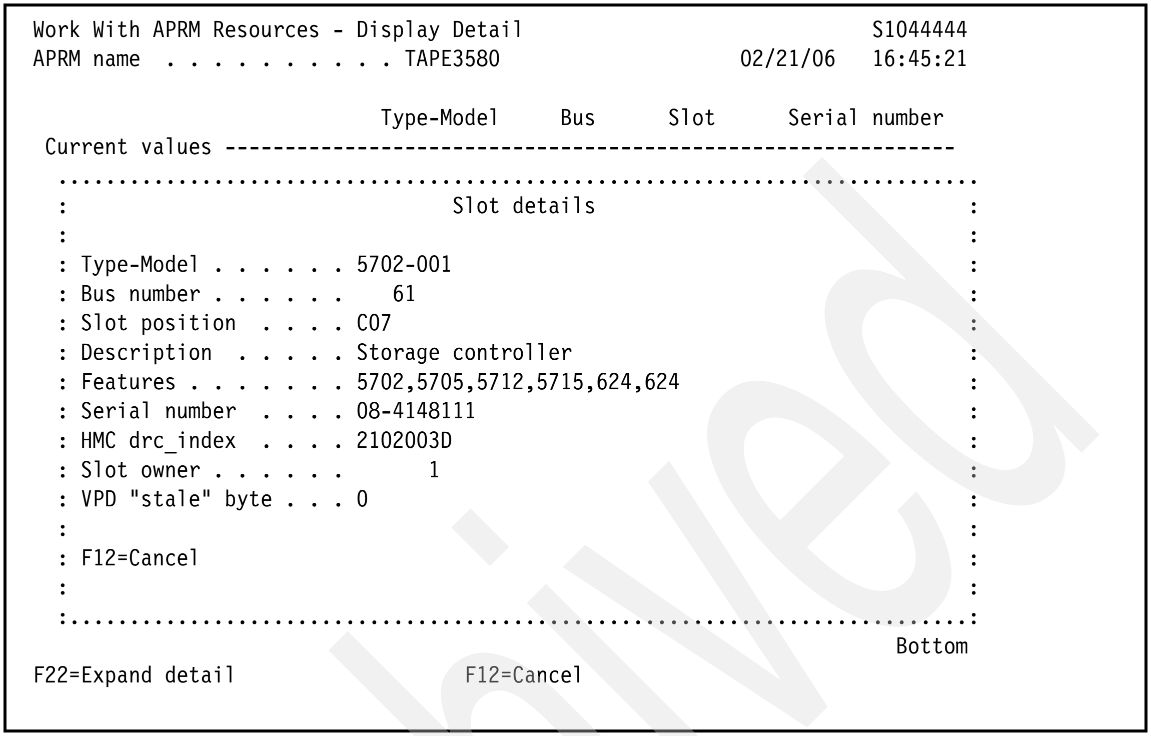


```
                   Work With APRM Resources - Display Detail          S1044444
APRM name  . . . . . . . . . . TAPE3580                    02/21/06   16:40:21

                             Type-Model     Bus      Slot      Serial number
 Current values --------------------------------------------------------------
 PCI I/O Processor            2844-001       61       C06       53-4083046
  Storage controller          5702-001       61       C07       08-4148111
  PCI 2-Line WAN w/Modem      2793-001       61       C08       42-33243F3




                                                                        Bottom
 Prior values ----------------------------------------------------------------
 PCI I/O Processor            2844-001       61       C06       53-4083046
  Storage controller          5702-001       61       C07       08-4148111
  PCI 2-Line WAN w/Modem      2793-001       61       C08       42-33243F3




                                                                        Bottom
F22=Expand detail                   F12=Cancel
```

Figure 8-3 Work With APRM Resources - Display Detail for System i

There is, in the System i version, a new command key that allows you to display more detail about any specific slot. If we choose slot `C07` in Figure 8-3 and press F22, we get Figure 8-4 on page 67.

Archived



```
Work With APRM Resources - Display Detail                             S1044444
APRM name  . . . . . . . . . . TAPE3580                    02/21/06   16:45:21

                             Type-Model     Bus      Slot      Serial number
 Current values --------------------------------------------------------------
  ..............................................................................
  :                               Slot details                                :
  :                                                                           :
  : Type-Model . . . . . . 5702-001                                           :
  : Bus number . . . . . .    61                                              :
  : Slot position  . . . . C07                                                :
  : Description  . . . . . Storage controller                                 :
  : Features . . . . . . . 5702,5705,5712,5715,624,624                        :
  : Serial number  . . . . 08-4148111                                         :
  : HMC drc_index  . . . . 2102003D                                           :
  : Slot owner . . . . . .       1                                            :
  : VPD "stale" byte . . . 0                                                  :
  :                                                                           :
  : F12=Cancel                                                                :
  :                                                                           :
  :............................................................................:
                                                                        Bottom
F22=Expand detail                   F12=Cancel
```

Figure 8-4 Work With APRM Resources - Slot details for System i

Archived

Archived

A

APRM command summary

Archived

After you install Automatic Partition Resource Manager (APRM), the following commands are available in both the product library (QBARSA) and the system library (QSYS):

- ACQAPRMDEV: Acquire APRM Device. This command allows you to automatically move input/output processors (IOPs) from one partition to another partition. You can execute this command from any OS/400 or i5/OS partition.
- CFGAPRM: Configure APRM. This command helps you complete the installation by specifying TCP/IP addresses, ports, message queues for information and exception messages, and user authorities to various levels of control and by synchronizing the passwords for the registered "APRM user." You can execute this command from any OS/400 or i5/OS partition.
- CHGAPRMTNR: Change APRM tuning rules. This command allows you to modify currently defined tuning rules for a partition. Use this command in the primary partition of an iSeries system or any i5/OS partition of a System i platform.
- CVTAPRMHST: Convert APRM History. Use this command in the primary partition of an iSeries system or any partition of a System i platform. This command extracts performance data from the APRM journal for the designated partitions, time period, and processing resources. The extracted data is averaged for each interval by partition and the results are placed in a

© Copyright IBM Corp. 2007. All rights reserved.

separate IFS file for each resource. Each record in these files contains the averages for the interval ending at the record's date and time.

- DSPAPRMDEV: Display Device's IOP location. This command, in the partition that it is run, finds the location of the IOP that controls the device and display or prints the bus, frame, and card position. It can be helpful for you to know this information before you decide which IOP to switch between partitions in order to time-share devices. You can execute this command from any OS/400 or i5/OS partition.
- ENDAPRM: End APRM. This command ends all the APRM jobs currently running in the partition where the command is executed. You do not typically need this command for daily operations because APRM usually ends by terminating subsystem QSYSWRK, typically during a system power down (PWRDWNSYS), by ending all subsystems (`ENDSBS *ALL`), or by executing the end system (ENDSYS) command. You can execute this command from any OS/400 or i5/OS partition.
- HLDAPRMTNR: Hold APRM tuning. Use this command to place a hold on processing resource tuning for one or all (*ALL) partitions. Use in the primary partition of an iSeries system or any i5/OS partition of a System i platform.
- MOVAPRMRSC: Move APRM Resource. Use this command in the primary partition of an iSeries system or any partition of a System i platform. This command allows you to change the ownership of one IOP and all the adapters and devices connected to it from whichever partition owns it to another partition (or to and from *AVAIL).
- RLSAPRMTNR: Release APRM tuning. Use this command to remove tuning holds that have been placed on one or all (*ALL) partitions. Use in the primary partition of an iSeries system or any i5/OS partition of a System i platform.
- RTVAPRMPTD: Retrieve partition details. Use this command to retrieve detailed data about a specified partition from APRM data files. Use in the primary partition of an iSeries system or any i5/OS partition of a System i platform.
- RTVAPRMPTN: Retrieve partition summary data. Use this command to retrieve summary data about the current partition from APRM data files. Use in the primary partition of an iSeries system or any i5/OS partition of a System i platform.
- RTVAPRMTNR: Retrieve APRM tuning rules. Use this command to retrieve detailed data about a specified partition's APRM tuning rules from APRM data files. Use in the primary partition of an iSeries system or any i5/OS partition of a System i platform.
- RUNHMCCMD**:** Run HMC command. Use this command on System i platforms only. This command allows you to execute any HMC command

Archived

through the APRM HMC interface. For example, to determine the code level of the HMC control program, you can execute the following command:

```
lshmc -V
```

This command might return:

```
“version= Version: 5
 Release: 1.0
HMC Build level 20051110.1
MH00507: Maintenance Package for V5R1.0 (12-03-2005)”
```

- STRAPRM : Start APRM. This command manually starts all necessary jobs to run the instance of APRM in the partition where it is executed. If some APRM jobs are already running, this command does not disturb them. You typically do not need this command for daily operations because APRM is usually started automatically during system startup procedures by an autostart job entry in subsystem QSYSWRK. You can execute this command from any OS/400 or i5/OS partition.
- SNDAPRMCFG: Send APRM Configuration. Use this command on System i platforms only. This command allows you to send all APRM configuration data, such as TCP/IP addresses and order of succession rules, from one partition to another partition. This command is especially useful when you create new partitions or one or more partitions have been inactive during a period when other APRM partitions might have had significant configuration data inserted. Data that is sent during the execution of this command includes TCP/IP addresses for all partitions and the HMCs, the order of succession rules, the IOP names for I/O hardware switching, tuning rules, and so on.
- WRKAPRM: Work with APRM. Use this command in the primary partition of an iSeries system or any partition of a System i platform. This command runs an OS/400-based utility (similar to the iSeries Navigator plug-in for APRM). You can view partition performance statistics, view and change tuning rules, and view machine configuration information.
- WRKAPRMRSC: Work with APRM Resources. Use this command in the primary partition of an iSeries system or any partition of a System i platform. You use this command to assign an “APRM name” to IOP resources. You can then switch these named resources to and from partitions or to and from available (*AVAIL - not assigned to any partition).
- Use the following debugging commands only when APRM Technical Support directs you to use them:
  - CHGAPRMDBG: Change APRM debug level. Use this command in the primary partition of an iSeries system or any partition of a System i platform. This command causes the partition Agent program, in any partition or all partitions, to log records to the APRMDEBUG data file, print reports, and clear that file.

Archived

  - CHGTNRDBG: Change TUNER debug level. Use this command in the primary partition of an iSeries system or any partition of a System i platform. This command causes the TUNER program to log records to the APRMTNRDBG data file and user entries to the QAPRM journal.

You can find detailed steps for requirements planning and LPAR implementation in the following IBM Redbooks publications:

- *Slicing the AS/400 with Logical Partitioning: A How-to Guide,* SG24-5439
- *Capacity Planning for Logical Partitioning on the IBM iSeries system,* SG24-6209

The IBM Redbooks publication Web site is:

http://www.redbooks.ibm.com

Archived

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this IBM Redbooks publication.

IBM Redbooks publications

For information about ordering these publications, see “How to get IBM Redbooks publications” on page 74. Note that some of the documents referenced here might be available in softcopy only.

- *IBM System i5 Handbook IBM i5/OS Version 5 Release 4,* SG24-7486

  `http://www.redbooks.ibm.com/redpieces/pdfs/sg247486.pdf`

- *IBM System i5, eServer i5, and iSeries Systems Builder IBM i5/OS Version 5 Release 4*, SG24-2155

  `http://www.redbooks.ibm.com/redbooks/pdfs/sg242155.pdf`

- *Logical Partitions on System i5: A Guide to Planning and Configuring LPAR with HMC on System i*, SG24-8000

  `http://www.redbooks.ibm.com/redbooks/pdfs/sg248000.pdf`

- *Slicing the AS/400 with Logical Partitioning: A How-to Guide,* SG24-5439

  `http://www.redbooks.ibm.com/redbooks/pdfs/sg245349.pdf`

- *Capacity Planning for Logical Partitioning on the IBM iSeries system,* SG24-6209

  `http://www.redbooks.ibm.com/redbooks/pdfs/sg246209.pdf`

Online resources

These Web sites and URLs are also relevant as further information sources:

- All APRM releases, “Highlights” documents, and technical manuals

  `http://www.barsaconsulting.com/aprmdownloads.htm`

Archived

© Copyright IBM Corp. 2007. All rights reserved.

How to get IBM Redbooks publications

You can search for, view, or download IBM Redbooks publications, IBM Redpapers, Hints and Tips, draft publications and Additional materials, as well as order hardcopy IBM Redbooks publications or CD-ROMs, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Archived

Index

Archived

© Copyright IBM Corp. 2007. All rights reserved.

Archived

IBM Redbooks

Automatic Partition Resource Manager for System i and iSeries

(0.2”spine)
0.17”<->0.473”
90<->249 pages

Archived

Archived

Archived

Automatic Partition Resource Manager for System i and iSeries

Understand the benefit of automating resource movement

Discover how to tune your system with APRM

See how APRM simplifies sharing I/O devices

Logical Partitioning for OS/400 opened the door to significant operational advantages and cost savings for iSeries and System i clients. You can now manage production workloads from many iSeries platforms on a single newer server using LPAR technology. And, in a System i world, you can now consolidate additional operating system environments onto a single System i footprint.

This IBM Redbooks publication is about setting up a partitioned iSeries or System i platform. The primary focus is about making it possible to share system resources between and among partitions easily, flexibly, and safely.

The contents of this book include information and a practical example of how to use an iSeries and System i tool, Automatic Partition Resource Manager (APRM®) from Barsa Consulting Group, LLC, to help transform a company's infrastructure into an on demand, self-healing environment.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-7174-00 ISBN 0738489298

Archived