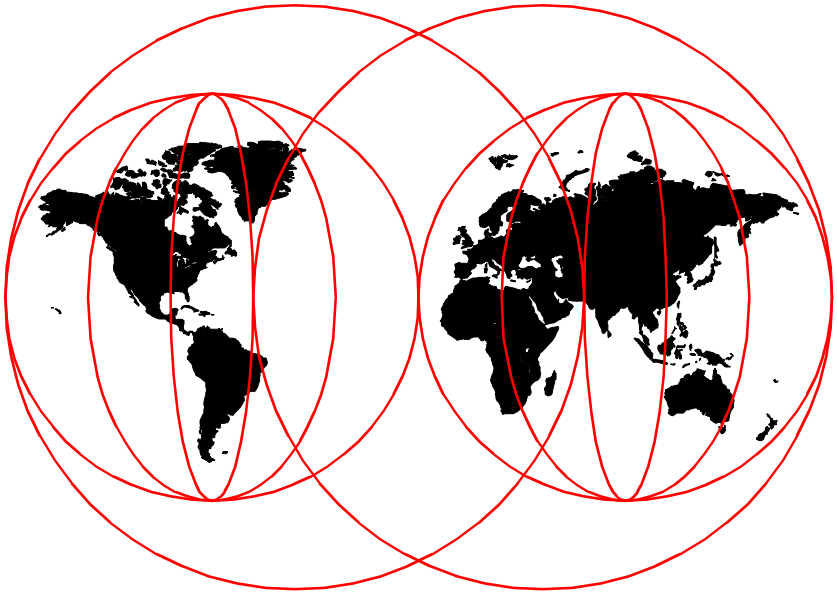
Slicing the AS/400 with Logical Partitioning: A How to Guide

Gottfried Schimunek, Danny Dupuche, Tim Fung, Paul Kirkdale,
Erik Myhra, Helmut Stein



International Technical Support Organization

http://www.redbooks.ibm.com

SG24-5439-00

SG24-5439-00

International Technical Support Organization

Slicing the AS/400 with Logical Partitioning: A How to Guide

August 1999

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix F, “Special Notices” on page 205.

First Edition (August 1999)

This edition applies to Version 4, Release Number 4 of the AS/400 Operating System OS/400, Program Number 5769-SS1

Note

This book is based on a pre-GA version of a product and may not apply when the product becomes generally available. We recommend that you consult the product documentation or follow-on versions of this redbook for more current information.

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. JLU Building 107-2
3605 Highway 52N
Rochester, Minnesota 55901-7829

When you send information to IBM, you grant IBM a non-exclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

Contents

Figures

Tables

Preface

Take a tour of logical partitioning on the AS/400! Logical partitioning allows one physical system to be divided into multiple logical partitions, each endowed with its own resources and capable of running independently. This redbook discusses the logical partitioning revolution on the AS/400 system. It is designed for all levels of understanding. It targets the person who simply needs an overall view of logical partitioning on the AS/400 system. And, it targets the technical people who will be actively engaged in devising solutions that involve logical partitioning, creating partitioned machines, and managing and operating them.

A shift to a more consolidated view of computer systems has triggered a renewed increase in machine sizes capable of consolidating multiple environments. The AS/400 system has embraced this trend by increasing its size and implementing logical partitioning, the concept of independent logical machines within large physical machines. This redbook is about the logical partitioning revolution on the AS/400 system. It is a new weapon in the AS/400 system's already rich armory.

Logical partitioning dates back to 1976, when IBM experimented with the S/370. Other hardware manufacturers followed this IBM path with various flavors of the concept. By 1997, companies such as Amdahl, Hitachi, and SUN, along with IBM, had their own versions of system partitioning. In 1999, the AS/400 system burst onto the scene with its own implementation.

Here is a brief guide to the chapters of this redbook:

- Chapter 1 explains the basic concepts of logical partitioning and provides several scenarios where this new functionality can be beneficial.
- Chapter 2 offers practical information for the reader who needs to plan and implement a logical planning solution. This chapter has numerous references to other existing publications for finer detail as required.
- Chapter 3 considers the aftermath of logical partitioning and addresses itself mainly to the day-to-day routine of working with logical partitioning.
- Chapters 4 and 5 extend Chapter 3 and deal with change management and problem management in the context of logical partitioning.
- Chapter 6 discusses capacity planning and performance. It provides valuable insight in the task of initially assessing the suitability of implementing logical partitioning in a given situation, developing the logical partitioning configuration to fit the situation, and ensuring the continued optimal performance of the installation.

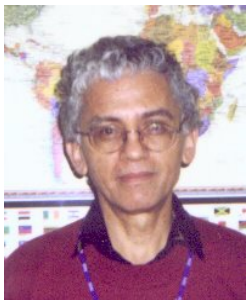
- Chapter 7 explains the crucial topic of high-speed communication between logical partitions. This is one of the many remarkable features of logical partitioning.
- The redbook contains appendices and references to other publications.

This material is also beneficial to IBM customers, IBM Business Partners, and IBM technical staff. Regardless of your level of expertise, this redbook tells you all you need to know about logical partitioning on the AS/400 system.

The Team That Wrote This Redbook

This redbook was produced by a team of specialists from around the world working at the International Technical Support Organization Rochester Center.

Gottfried Schimunek is a certified IT Architect at the IBM International Technical Support Organization, Rochester Center. He writes extensively and teaches IBM classes worldwide on all areas of application development and performance. Before joining the ITSO in 1997, Gottfried worked in the Technical Support Center in Germany as a consultant for IBM Business Partners and customers as well as for the IBM sales force.



Danny Dupuche is a Senior Technical Consultant with KAZ Computers Services, a major IBM Business Partner in Australia. Originally from a mainframe background, Danny started working with the S/38 in the early years of that system. His areas of expertise are in capacity planning and sizing, systems management, and high availability solutions for the company's customers. He currently monitors and analyzes customer business requirements and, in conjunction with the company's marketing arm, provides solutions to meet those goals.

Tim Fung is a Consulting I/T Specialist at IBM China/Hong Kong Limited. He has over ten years of experience in the S/36, S/38 and AS/400. He works with the AS/400 in IBM Hong Kong providing technical marketing support to AS/400 sales representatives, IBM Business Partners and customers. His areas of expertise include availability solutions, capacity planning, and the Internet.

Paul Kirkdale is a certified AS/400 IT Consultant working for IBM in the UK. He has 10 years of experince in working with the AS/400. Currently, he provides technical support for high availability and clustering to large AS/400 users in the UK and across Europe. Prior to this role, he worked in the areas of application development, database design, performance, and systems management. Before joining IBM in 1990, Paul worked for an IBM Business Partner where he provided AS/400 and System/38 support to customers.

Erik Myrha is a Senior Advisory SE at the Customer Support Center in Norway. His areas of expertise include application development and performance. He has worked at IBM for 22 years, including three years in the Rochester ITSO.

Helmut Stein is an AS/400 IT Consultant working for IBM in Germany. He has 23 years of experience in the S/34, S/36, S/38, and AS/400 systems. He provides technical pre-sales support to large AS/400 customers, IBM sales representatives, and IBM Business Partners. His areas of expertise include performance, high-availability, server consolidation, and OptiConnect.

A special thanks goes to **Joel Nelson** from the Rochester Product Field Support Center. He was very instrumental in finding answers related to

AS/400 hardware configurations and the use of Hardware Service Manager, as well as the dedicated service tools.

Thanks goes also to the following people for their invaluable contributions to this project:

Doug Anderson
Bill Armstrong
Sue Baker
Dan Birkestrand
Lynn Chung
Amit Dave
Selwyn Dickey
Mike Faunce
Ian Jarman
Mike Leslie
Lynn McMahon
Mark Manges
Naresh Nayar
John Reichel
Debbie Saugen
Jeff Scheel
Brenda Thompson
Garrett Winn

Comments Welcome

Your comments are important to us!

We want our redbooks to be as helpful as possible. Please send us your comments about this or other redbooks in one of the following ways:

- Fax the evaluation form found in “ITSO Redbook Evaluation” on page 219 to the fax number shown on the form.
- Use the online evaluation form found at: `http://www.redbooks.ibm.com/`
- Send your comments in an Internet note to: `redbook@us.ibm.com`

Chapter 1. Logical Partitioning Concepts

This chapter explains the basic concepts of AS/400 Logical Partitioning. The end objective of logical partitioning (LPAR), released with OS/400 V4R4, is to provide users with the ability to split a single AS/400 system into several independent systems capable of running applications in multiple, independent environments simultaneously. For example, logical partitioning makes it possible for a user to run a single application using different sets of data on separate partitions, as if it was running independently on separate physical AS/400 systems.

1.1 Logical Partitioning History

IBM began the study of logical partitioning for the S/370 architecture in Poughkeepsie, New York, in 1976. The project proved that logical partitioning was a workable concept, but the idea had not been extended to actual product implementation. In November 1984, Amdahl Corporation announced the 580/Multiple Domain Feature on the IBM S/370 mainframe. It was the first logical partitioning product and it was favorably accepted by the market. Then, the evolution of partitioning began (see Figure 1).

Date	Vendor	Product Name
1985	Amdahl	Multiple Domain Feature on IBM Mainframe
1988	IBM	PR/SM Logical Partitions on IBM Mainframe
1989	Hitachi	Multiple Logical Processor Facility on IBM Mainframe
1997	SUN	Dynamic System Domains on SUN UE 10000 server
1999	IBM	Logical Partitioning on AS/400

Figure 1. Chronology of Logical Partitioning Implementation

As we can see, partitioning has been available on IBM mainframes for more than a decade. Over that period, it has evolved from a predominantly physical partitioning scheme, based on hardware boundaries, to one that allows for

virtual and shared resources, with dynamic load balancing. In today's marketplace, partitioning has become a requirement. All the major mainframe players, including IBM, offer partitioning.

History repeats itself. The factors that have driven the evolution of mainframe partitioning over the past decade are now at work in the server system arena. Partitioning is fast becoming a necessity there too. It is estimated that by the year 2001, all major players in the server marketplace may offer some degree of partitioning.

The AS/400 system is delivering its own complete version of partitioning, right now. Logical partitioning implementation on an AS/400 system is an adaptation of the System 390 logical partitions concept, with flexible and granular allocation of system resources. It also offers flexibility in allocating interactive performance and high-speed internal communications between partitions.

The current hardware offering delivers a maximum of 12 processors, which allows a maximum of 12 logical partitions. The same standard tools used across physical machines, for example SQL, DDM, DRDA, Remote journaling, and DB2 Multisystem, are available for use between logical partitions.

1.2 Scenarios for Logical Partitioning

The range of scenarios possible with logical partitioning is limited only by your imagination. Let us select a few of the most promising ones and explore the solutions offered by the application of logical partitioning to see what can be achieved.

1.2.1 Multiple Production Environments

Let us first consider the case of a company that offers outsourcing services to its customers. The economics of the case convinced the CEO that the consolidation of all the company's physical machines into a single multipartitioned machine would deliver improved services to customers and streamline day-to-day operations. The profit expectations look attractive.

This service provider can purchase a multiprocessor model 7XX with enough resources to create several completely independent logical partitions to meet customer requirements. These partitions can operate as independent machines. For example, some partitions may run client/server workloads, while others run purely interactive workloads, still others may run mixed workloads. The permutations are endless.

1.2.2 Mixed Production and Test Environments

Generally, production and test environments must be kept separate. Logical partitioning achieves this objective in a simple way.

One or several partitions can be allocated for production only. Others are set up to handle application testing, either to validate application changes or to test applications under new releases of system software.

One special case today is testing applications for Year 2000 ready (Y2K) compliance. A logical partition can be set up with its own special environment having its own date, time, and system values. Communications between the partitions can be set up as required. Figure 2 illustrates this concept.

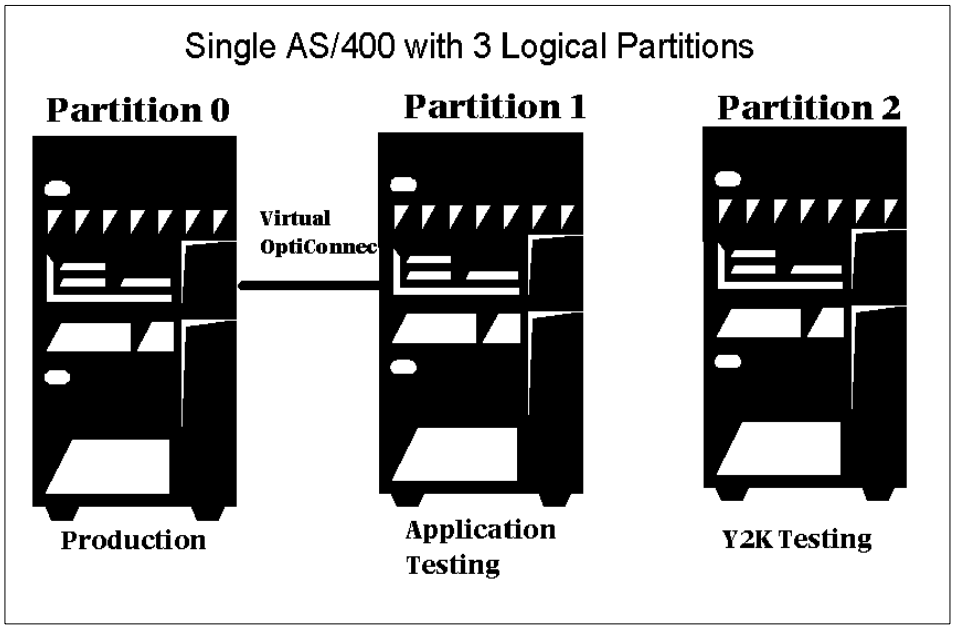


Figure 2. Mixed Production and Test Environments

Without logical partitioning, the only practical way of testing Y2K would be to purchase additional hardware and software. However, the disadvantage may be that the hardware and software may not be required subsequently. Note that system software is currently charged at the physical machine level. Therefore, it can be run on several independent partitions, without an additional charge.

Partitioning provides a solution by making it possible to use an independent partition for testing. When testing is over, the resources allocated to this partition can be returned to the production partition or elsewhere as required.

1.2.3 Implementing an Archiving Solution

In this scenario, the customer has a very large database containing a mix of live and historical data. Most of the interactive and batch jobs use the live records on the database. However, the customer needs to access the historical data both interactively and in batch mode from time to time, with reasonable response times. Furthermore, they also need to reduce their backup time, without jeopardizing data security. The parameters of this situation are:

- Database size.....500GB
- Number of logical views..... 60
- Live data..... 200GB
- Historical data..... 300GB

Assuming that the application can distinguish between live and historical data, one possible solution is to create a three-partitioned configuration as shown in Figure 3.

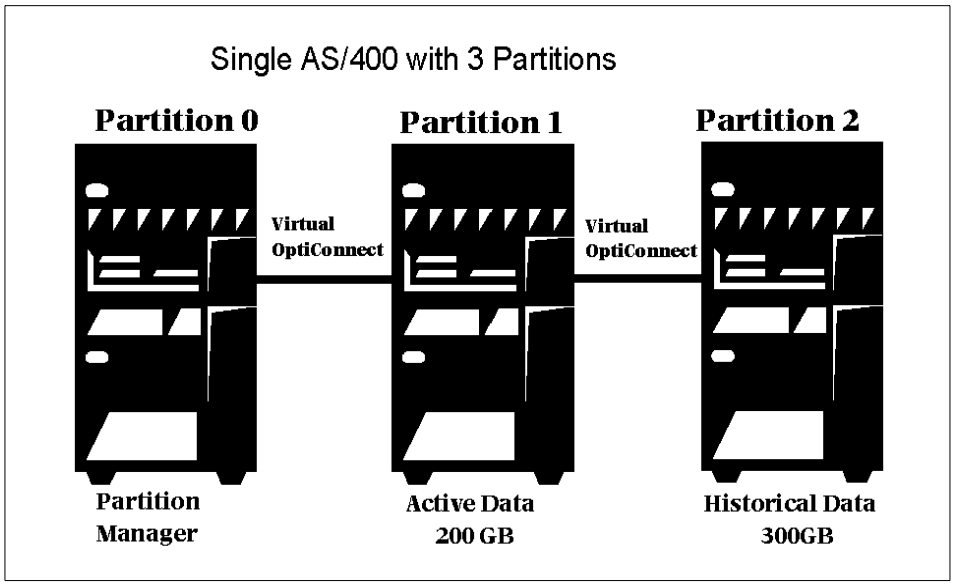


Figure 3. Setting Up an Archiving Solution

The outcome of this implementation would be:

- Faster updates of live data because of reduced database index tables.
- Seamless access to historical data through high-speed inter-partition communication.
- Reducing the size of the database also reduces the size of the database index tables and, therefore, the amount of work the CPU has to do to update these tables.
- Overall improved response times.
- Reduced backup requirements. Only live data needs to be backed up at frequent intervals. Backing up archived data needs to be performed only at the time when live data transitions to archived status.
- Report consolidation through DB2 Multisystem or Remote SQL over high-speed inter-partition communication.

1.2.4 System Consolidation

Today's AS/400e server offers over 300 times the performance range from the smallest to the largest servers. Many companies are taking advantage of the high degree of scalability by consolidating existing AS/400e applications onto fewer systems, often in regional data centers.

Still many companies are attracted to the concept of distributed computing, in particular on the AS/400, Unix, and NT servers. The attraction of distributed servers is due to the low initial investment costs and the flexibility of deploying applications customized to the needs of specific departments. However, the realities of managing distributed UNIX and NT platforms proved far more difficult and costly than initially imagined.

In response, many companies are now investigating opportunities to consolidate their decentralized and distributed systems to regain both financial and administrative control of their IT resources. This trend in the industry is referred to as server consolidation. In a recent International Data Corporation (IDC) white paper, *Server Consolidation: An AS/400 TCO Analysis*, 49% of respondents to their survey were in the process of consolidating servers and storage.

With its strengths as both an enterprise and distributed server, the AS/400 system has a unique position as a server consolidation platform for AS/400 consolidation, Domino for AS/400, and Windows NT Integrated Netfinity Server. With the support of logical partitions on the AS/400 system, server consolidation becomes a key element of the AS/400 enterprise systems future.

IBM has a worldwide server consolidation solutions team to assist customers that are considering server consolidation. IBM has a range of consolidation solutions. They include a consistent methodology and a wide range of services, from expert advice in planning to optimizing and benchmarking consolidated applications. In addition to offering planning services, IBM Global Services has experienced specialists who can perform an AS/400 consolidation.

For more information, see IBM's Web site at: `http://www.as400.ibm.com/sc`

This Web site includes a wide range of information on server consolidation, including reference to the following consultant reports:

- *Getting It Together: AS/400e — New Ammunition for Server Consolidation* (D.H Andrews Group)
- *Server Consolidation: An AS/400 TCO Analysis* (International Data Corporation)

1.2.5 Multidatabase Applications

Some businesses may require running the same application against independent databases. It may be that, because of naming conventions or other internal restrictions, the application cannot be run against multiple independent databases on the same physical machine. The solution before logical partitioning would be either to purchase additional hardware or modify the application.

With logical partitioning, each instance of the application can be run within its own independent partition. In addition, communications between the independent partitions can be used for consolidated reporting if required.

1.2.6 Minimizing Backup and Recovery Windows

The imperatives driving today's businesses demand total machine availability, 7 days-a-week, 24 hours-a-day (7 by 24). On the other hand, data security demands that a foolproof backup strategy be in place to meet unforeseen contingencies. A third ingredient in this situation is the relentless growth of databases as businesses become more complex. Logical partitioning can provide one solution to balance these conflicting needs.

Let us look at a scenario where backup is becoming so time consuming that the production window is dwindling fast. Figure 4 illustrates a possible solution to minimize backup of the production database. Incidentally, this scenario also facilitates recovery.

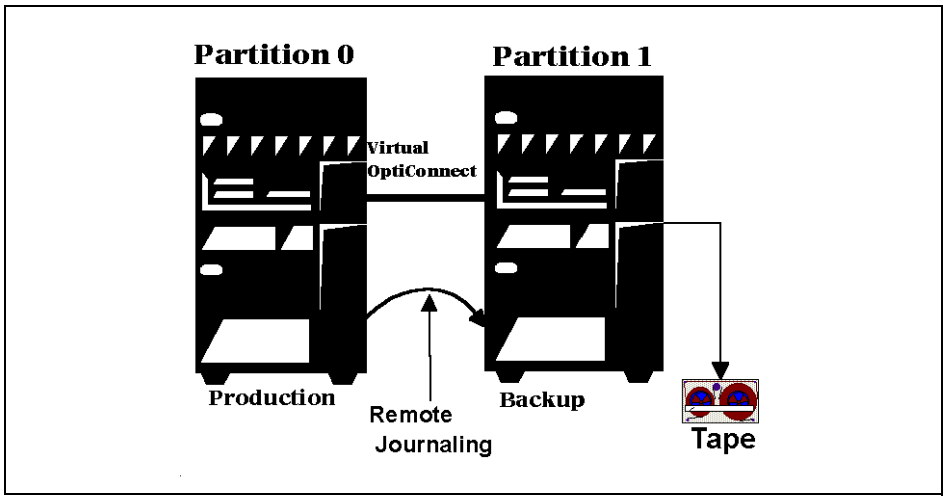


Figure 4. Minimizing Backup

In this scenario, production is running on partition 0 and all updates are replicated on partition 1, using remote journaling. At preset intervals, the partition 1 update is suspended, and the partition database is backed up. After the backup, the partition 1 database is re-synchronized with partition 0, by applying all accumulated journal entries from partition 0 to partition 1.

This scenario provides for recovery of the production database onto the same partition or a different physical machine, with minimum inconvenience.

1.2.7 Consolidating Worldwide Operations

Time zones are a way of organizing time so that people living in a large area (country or part of a continent) share the same time, just to make things easy. People living at about the same longitude usually have about the same local time. There are 24 standard time zones around the world. These are all an integral offset of hours to Universal Time Coordinated (UTC)-time, for example, +5 hours or -10 hours to UTC. If UTC-time is 12:00 (AM), then people having offset +5 hours, have local time 17:00 (or 5:00 PM).

If your business needs different time zones on a single AS/400 system, there are no AS/400 system values available to support you. If you, for example, have a system located in Hong Kong and have offices located in the US, the users signing on to that system from the US offices will always have the Hong Kong time. Your application has to manage the different time zones for your users, which may be located around the whole world (Figure 5). Even if you can manage the different time zones by your application, you still need to find

a way to manage your downtimes. Your system cannot be saved while your users in the US are still running their interactive applications. You need to have batch windows for running batch programs.

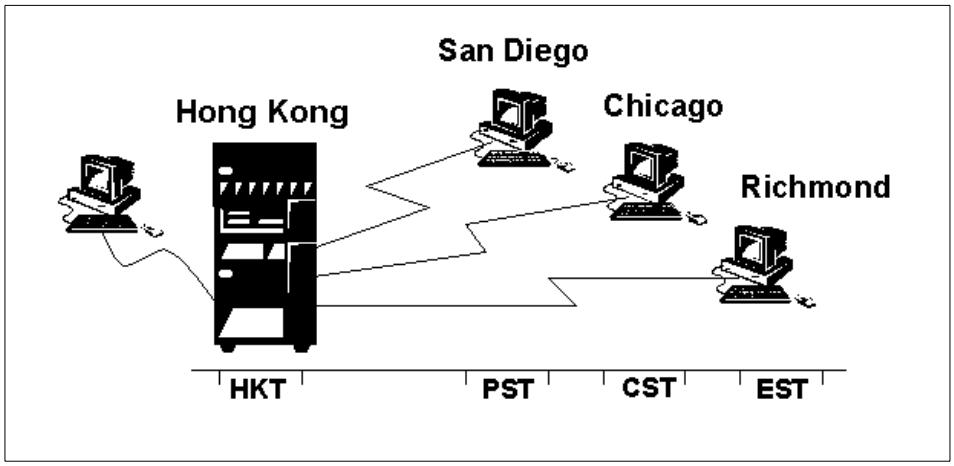


Figure 5. Different Time Zones on One AS/400 System

One way to provide different time zones to your users is to split your AS/400 system from a single system image to a system with logical partitions. Even if the time is only provided by the service processor of the primary partition, you can have different times on logical partitions. The logical partitions manage the time difference to the primary partition.

Usually your application has only one database for all of your central and distributed offices. You have to decide whether you want a database partition and application partitions for all your different time zones or you want your central site application and database in one partition.

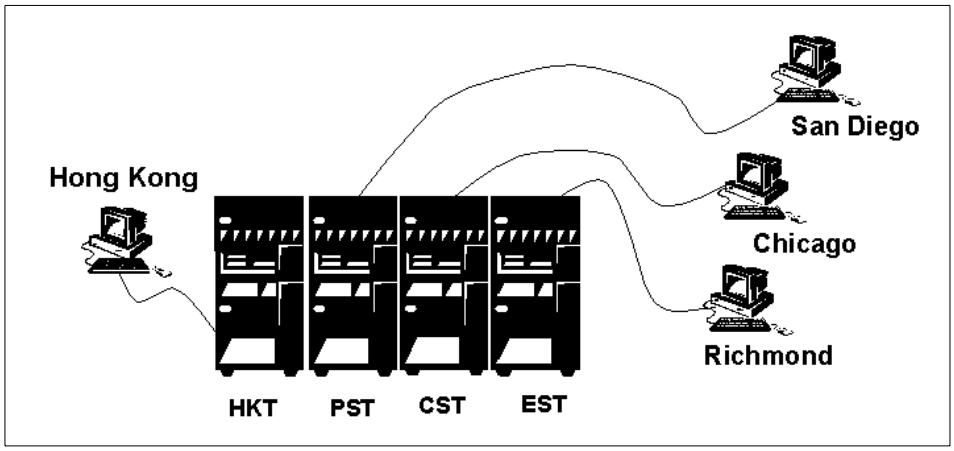


Figure 6. Logical Partitions for Different Time Zones

In Figure 6, the single AS/400 system image was split up into four partitions. In the primary partition, you have the database for all secondary partitions. For each time zone, you have your own logical partition. The application in the logical partition has to be changed. Your database resides in the primary partition, and the application in the logical partition needs to know where the database is.

Using this method for different time zones on a partitioned system, you still need to manage your downtime. There is still only one image of your database, and you have to stop your applications while you backup your database. If you need to provide 24-hour service for your users, you need to mirror your database on a logical partition. Different software vendors provide solutions for doing that. The backup is done only from the secondary partition, and your users can work with the applications for 24 hours.

1.2.8 A Perfect Combination with Domino Clustering

E-mail and groupware software, such as Lotus Domino for AS/400, becomes more and more critical in the daily business operation. Customers cannot afford to have their Domino servers fail for hours. This may affect customer service levels.

Lotus Domino for AS/400 has already taken advantage of the reliability and availability features of the AS/400 system, such as RAID-5, mirrored disk units, and integrated backup capability. With the AS/400 system's logical partitioning and Domino clustering, the level of availability is further enhanced.

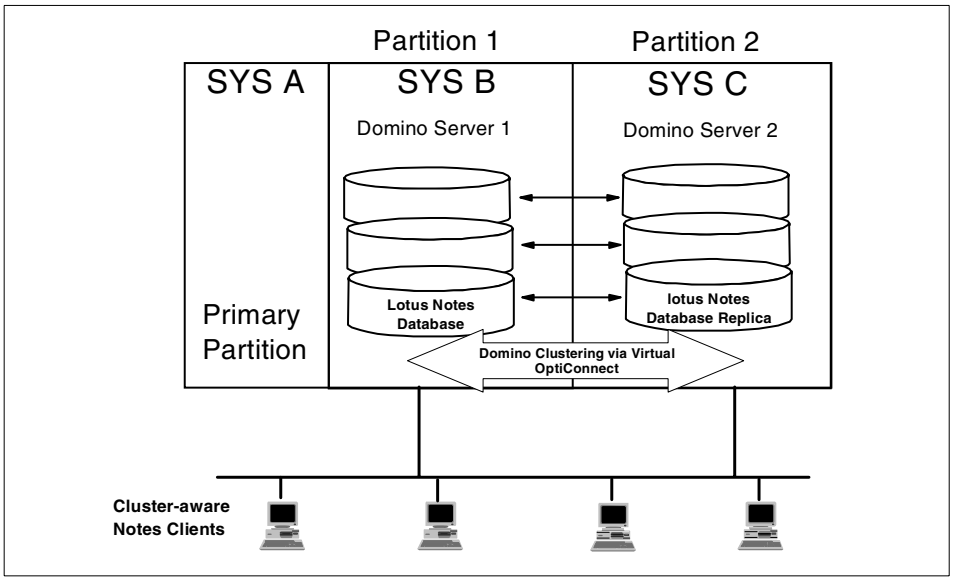


Figure 7. Lotus Domino Clustering over AS/400 Logical Partitions

The cluster-aware Notes clients in Figure 7 use Domino server 1 as the home server. When the Notes client tries to open a database on the Domino server 1 and the server is down, the clients look for the next available server on the server list in the client's own cluster cache. Then, they access the Cluster Manager on Domino server 2 in the cluster. The cluster manager then redirects the client to Domino server 2. The user continues to have the database access on server 2. The cluster replication ensures that all the changes to the database are immediately passed to the database replica in the cluster. Databases are tightly synchronized to provide the high availability.

Putting the Domino servers into two different AS/400 logical partitions isolates them from software errors affecting both servers. Plus, it also provides the ability to have high-speed cluster replication between the two AS/400 partitions using the virtual OptiConnect function.

1.2.9 Three-Tiered Application Architecture

Several types of AS/400 applications use a three-tiered architecture. For example. Enterprise resource planning (ERP) applications often use a desktop PC, an application server and a back-end database server. With the availability of AS/400 logical partitioning, customers can put the database server in the primary partition and application servers in secondary partitions within the same system. This can reduce footprint and maintenance costs.

Furthermore, the database server can communicate with the application servers through the virtual OptiConnect function available between the partitions.

1.3 How It Works

Logical partitioning (Figure 8) is achieved by distributing the resources of a single AS/400 system to create multiple independent systems, within the same system. The resulting structure consists of a primary partition, and one or more secondary partitions.

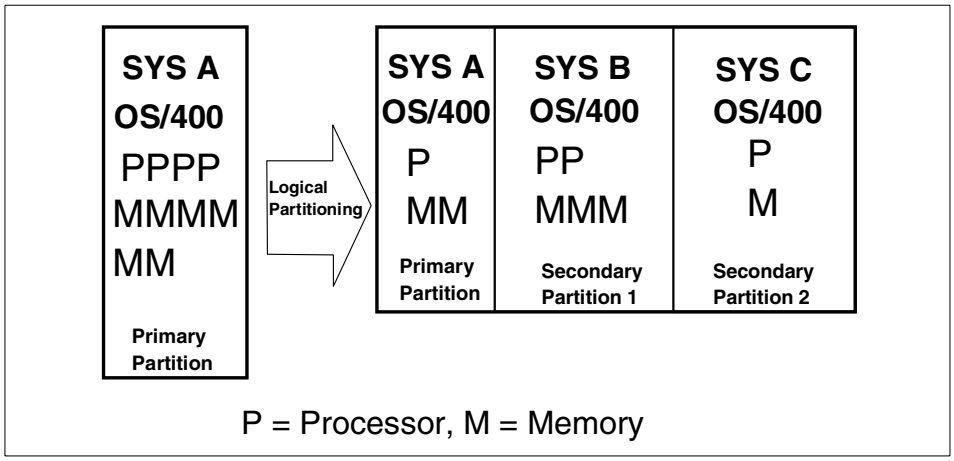


Figure 8. AS/400 Logical Partitioning

Resources, such as processors, memory, and I/O devices, are logically assigned to primary and secondary partitions.

1.4 Partitioning Terminology

The new nature of logical partitioning demands that all terms in use be clearly understood. Let us review the hardware and software components that we will use to create and manage logical partitioning, and define some new terms.

This book is structured to offer you the flexibility of browsing. At this point, you may want to refer to Appendix A, “Logical Partitioning Terminology” on page 167. This appendix includes the terms that we are going to use in the process of working with logical partitioning. You may also want to refer to Figure 9 on page 13 to see where everything fits into the context.

1.5 Primary Partition

When OS/400 is installed on a new system, it is always configured with the primary partition only. This partition owns all the resources available on the machine, such as processors, main storage, and system buses. It functions as one of the logical systems and also provides management functions for the configuration of the secondary partitions, such as:

- Power Management (includes concurrent maintenance)
- Virtual Operations Panel (controls power up and down, IPL, and virtual key lock)
- Logical Partition definition (for configuring logical partitions)

The primary partition can also function as a hub for external communications or Electronic Customer Support for the entire system.

1.6 Secondary Partitions

Secondary partitions are created and managed from the primary partition, but function as independent systems within the physical system. They have their own resources such as processors, main storage, and system buses. Plus, they have their own primary language, system values, time-of-day, user profiles, OS/400 system Licensed Internal Code (SLIC), IBM Licensed Programs (LPP), database files, and user applications. Furthermore, they can be independently powered down and up without affecting the primary or other secondary partitions. However, they retain some dependencies on the primary partition. For example, performing a system restart or power down on the primary partition powers down all secondary partitions.

This means that, before such actions are undertaken, all secondary partitions must be powered down to avoid possible abnormal secondary partitions termination.

1.7 Conceptual View of Logical Partitioning

Once you have reviewed the most common terms related to logical partitioning, we can see how it all fits together in Figure 9.

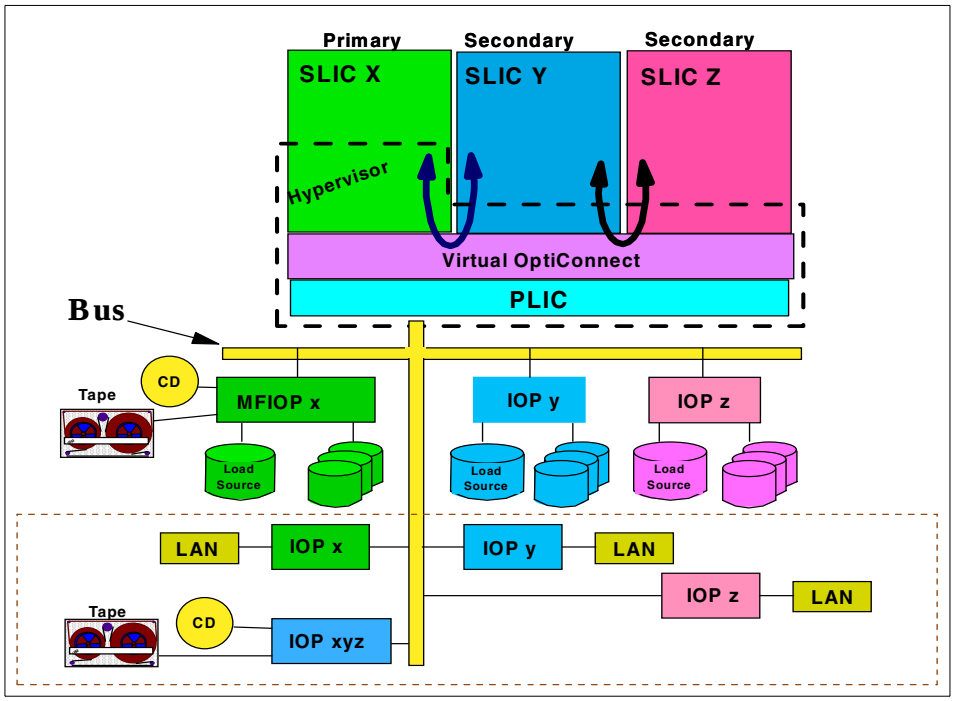


Figure 9. Generic View of Logical Partitioning

Figure 9 shows how a single system can be configured to have three partitions, with one as the primary partition. The logical partitions are also interconnected using the AS/400 system bus architecture through virtual OptiConnect to achieve high-speed data transfer between the partitioned systems. Each partition shows some dedicated I/O resources, such as disk drives, communications links, and workstation controllers. The example also shows additional resources, such as tape drives, CD-ROM drives, and communications adapters that can be dynamically shared by all of the partitions.

1.8 Clustering

Logical partitioning creates independent systems within a single physical box. Clustering can be seen as a superset of logical partitioning in that it provides a single resource view that binds together two or more physical AS/400 systems. These can, in turn, be logically partitioned if required.

Clustering offers increased high availability of applications and data. It does this by providing automated recovery facilities across several combined AS/400 systems (defined as nodes in the cluster). For example, through clustering, an application can be set up to run on a primary node, with a secondary node defined for quick fail-over switching, when the primary node becomes unavailable. The switch-over would include the automatic switching of communications to the secondary node.

It is also possible to imagine a cluster containing logically partitioned nodes. An application can run in one partition, in its own primary node, and have its backup secondary node ready on a logical partition in another node, somewhere in the cluster.

Figure 10 illustrates the concepts of clustering.

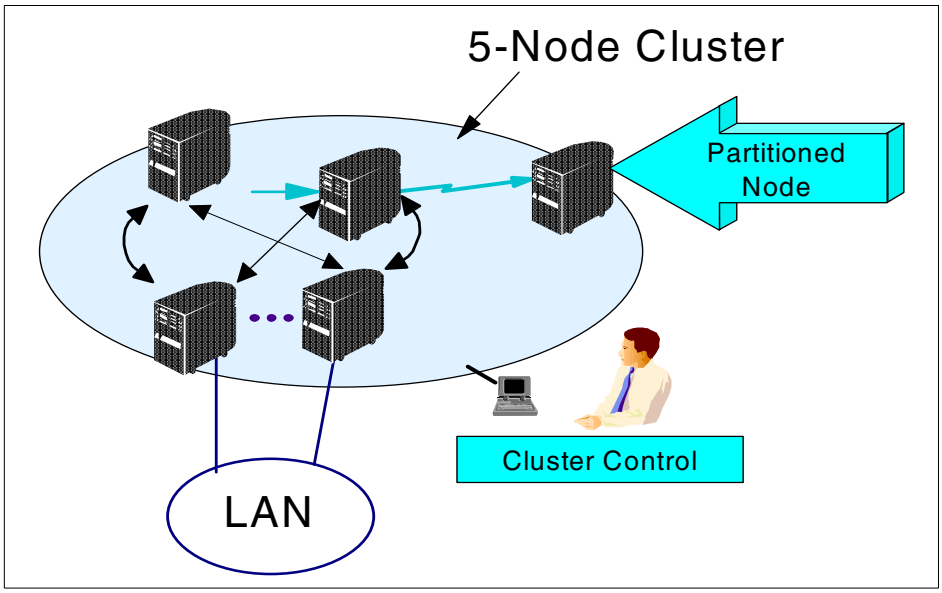


Figure 10. Clustered Configuration with a Partitioned Node

1.9 Marketplace Positioning

As we mentioned at the beginning of this chapter, system partitioning has now become a must in the server marketplace. It is interesting to look at some of the active players that offer partitioning solutions (see Table 1).

Table 1. Logical Partitioning Marketplace Offerings as of February 1999

	S/390	AS/400e	SUN*	HP 9000*
Operating Systems	All	OS/400	Solaris	HP-UX
Supported Servers	All Enterprise Servers	6xx, Sxx, 7xx	E10000	None
Partitioning Type	Logical	Logical	Physical	None
Maximum # partitions	15	12	8	
Maximum # processors / partition	12	12	64	
Minimum # processors / partition	<1	1	1-4	
Processor Increment above minimum	<1	1	1-4	
Memory Increment above minimum	1-4 MB	1MB	512 MB	
I/O Increment above minimum	Device	IOP	2 Buses	
Independent Movement of Resources	Yes	Yes	No	
Dynamically Change Resources	Yes for MVS	Planned, Stage 2	Yes database restart required to use	
Partition Weights	Yes	Interactive	No	
High speed internal communications	No	Yes	No	

1.10 What Makes the Difference

In 1.2, "Scenarios for Logical Partitioning" on page 2, we discuss some of the ways in which logical partitioning can be beneficial. The reality is that the permutations are limited only by your imagination and circumstances.

The AS/400 system now offers many more options that can be manipulated to meet every possible customer requirement involving physical machine consolidation and logical machine independence. The entire offering is packaged using the best technology available in the market place. To top it all, the AS/400 system product line has a wide and open-ended range that is capable of satisfying from the simplest to the most complex customer environments imaginable. Keep reading the remaining chapters of this book to discover how you can put it all together.

Chapter 2. Planning Considerations

Detailed planning for logical partitions is crucial. The effort made during the planning stages makes the transition to a partitioned environment a smoother process. Insufficient planning results in problems during the configuration of the new environment.

This chapter summarizes the areas you need to consider when planning for logical partitions. It is not intended as a replacement to the formal planning process that is documented on the logical partitioning home page at:
`http://www.as400.ibm.com/lpar`

This Web site contains the latest information available and provides links to these three essential articles:

- *Planning for and Setting Up Logical Partitions*
- *Backing Up and Recovering Logical Partitions*
- *Managing Logical Partitions*

2.1 Understanding Customer Requirements

Logical partitioning is a new mode of operation for the AS/400 system where multiple copies of OS/400 run on a single physical set of hardware.

Physically, the system will be packaged in exactly the same way that it is today. Multiple racks contain various types of I/O devices. However, each system has only one main system rack that contains processors, memory, and the system control panel.

Logical partitions can be considered a suitable solution for many different requirements. In some circumstances, partitions only meet some of the customer's overall requirements. The requirements that are not met need to be identified. The impact of this to the business should be evaluated.

2.1.1 Single Points of Failure

A machine that is configured for logical partitions has points of failure that can affect all partitions on that physical machine. A processor or memory card failure in any partition can cause the entire machine to end abnormally. Then, each partition has to perform an IPL to recover.

Any software or hardware problems that cause the primary partition to fail also cause all the secondary partitions to fail.

2.1.2 Availability Requirements

When using logical partitions for AS/400 hardware or application consolidation, the planning must balance the optimum consolidation possible with the customer's availability requirements.

It is possible to have a scenario where data from one partition is replicated to another partition on a different AS/400 system. This can be done using one of these options:

- High Availability Business Partner (HABP) applications
- The remote journalling function of OS/400 with a user defined process to apply the journal entries on the target partition
- A customer-written application

All three approaches can use either traditional LAN/WAN communications or the Opticonnect/400 function discussed in Chapter 7, "Inter-Partition Communication" on page 133.

Each AS/400 system then acts as a partial standby machine to the other. In the case of a software failure or specific hardware failure, only the workload from that partition has to be moved to the backup partition on the second physical machine.

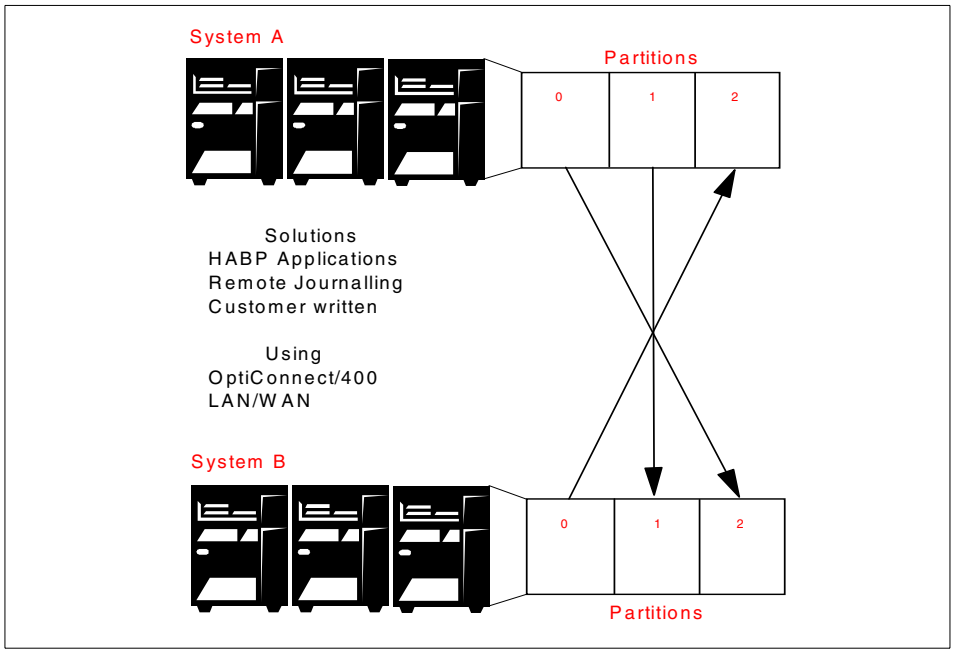


Figure 11. Replicating Data from One Partition to Another

Applications that have vastly different availability requirements may not work well together on the same AS/400 system.

2.1.3 Resource Requirements

Applications that have opposing peak resource times would fit well together on the same machine. Consider the example where application A has a high CPU requirement during the online day but very little overnight. This can be consolidated with an application that has very little CPU requirement during the day but a high requirement overnight. Appropriate planning is required to factor in the time needed to move CPU and memory resources between partitions.

2.1.4 Application Stability

The stability of a customer's applications may impact the selection of applications that run in separate partitions on the same AS/400 system.

An application that is fairly new into production may have yet to reach an adequate level of stability. This instability may impact the operational requirements and level of day-to-day management required.

Applications that have similar upgrade and maintenance requirements are better suited on the same AS/400 system.

2.1.5 Security Requirements

Every application and every system has its own security requirements. Appendix C, “Accessing LPAR Functions” on page 175, provides a list of all partioning functions and indicates from where they can be accessed. Consider this information when selecting applications to run in different partitions on the same AS/400 system.

2.2 Understanding the New Complexity of Your Environment

It is important to understand the new operating environment that is created as a result of setting up partitions. New functions, such as the fast inter partition communication discussed in Chapter 7, “Inter-Partition Communication” on page 133, open up more options to developers and support staff alike. These new functions, plus the inter partition dependencies for IPL and maintenance lead to new complexities for operational staff. The following sections describe some of the complexities of each management discipline.

Operations Management
Operators are introduced to new terminology and possibly new tasks to perform from unfamiliar menus. Technical support staff must be aware of the new environment and the fact that a change in one partition may impact another partition. For example, with the current implementation, allocating more memory to one partition requires it first to be removed from another.

Problem Management
Help desk staff must be aware of the new environment so they can perform comprehensive problem determination. The performance for inter partition communication is far better than what is available today using external LAN/WAN communications. Applications that read and update data between partitions are likely to use these new facilities.

Change Management
Application changes may need to be implemented a different way. See 4.2, “Managing User Applications in a Logical Partition” on page 63, for more details.

Performance Management
Performance and capacity planning need to take into account partition processing as well as application performance. See Chapter 6, “Capacity Planning and Performance” on page 81, for more details.

Configuration Management
Any system configuration changes that are made are likely to impact all partitions. Extra planning is required when trying to implement a change to the system. The dynamics of managing multiple systems have not changed with logical partitioning. You still need to view the system management disciplines across a logically partitioned system in exactly the same manner as you would do with multiple AS/400 systems.

2.3 Logical Partitioning Configuration and Ordering Process

The following steps outline the process to be followed when IBM or a business partner wishes to configure and implement logical partitions for a customer. The actual process may vary from customer to customer.

1. Find out more on logical partitioning. For further information, refer to the logical partition home page at: `http://www.as400.ibm.com/lpar`
2. Go to the logical partition home page, and print and review the three essential articles:
   - *Planning for and Setting Up Logical Partitions*
   - *Backing Up and Recovering Logical Partitions*
   - *Managing Logical Partitions*
3. Go to the logical partition home page (see step 1), and review the *AS/400 Logical Partitions Hardware Planning Guide*.
4. Perform a capacity plan from a logical partitioning perspective.

   The normal capacity planning process needs to take into account the performance available from the resources allocated to a partition. For more information, refer to Chapter 6, “Capacity Planning and Performance” on page 81.
5. Complete the Logical Partition Planning Worksheet.

   The planning worksheet is a framework for providing the information necessary to build an AS/400 system capable of supporting partitions. This can be found as a link from the logical partition home page and should be printed and then completed.
6. If partitioning is to be implemented on an existing AS/400 system, print an accurate hardware rack configuration. Follow these steps:
   a. Enter the `STRSST` command.
   b. Select option **1** (Start a Service Tool).
   c. Select option **7** (Hardware Service Manager).
   d. Press **F6** (Print configuration).

If partitioning is to be implemented on a new AS/400 system, print the configuration report from the PC configurator.

7. Contact the Rochester Technology Solutions Center (TSC) and supply them with the completed paperwork from step 4 and 5.

   They can be contacted by e-mail at: `rchtsc@us.ibm.com` or by fax at: 1-507-253-0424.

   The IBM Representative or Business Partner will arrange for a technical conference call between all parties if this is required.

8. Conduct a Systems Assurance review with the IBM Representative, Customer and Business Partner.

9. Order the necessary hardware and software (V4R4M0).

   The TSC will return a list of additional hardware that needs to be ordered to support the planned partitioning environment.

10. Install V4R4M0.

    We recommend that V4R4M0 is installed and run for one complete business cycle before any other changes are made to the system. This allows V4R4M0 to stabilize in the customer environment.

11. Install the new hardware.

    The TSC will provide the necessary card placement information necessary to support the specific partition requirements. This may require moving some existing hardware from their current card positions.

12. Set up the logical partitions.

    Do this by following the process described in using the *Planning for and Setting Up Logical Partitions* article.

13. Verify that all the hardware resources are reporting correctly in each partition.

    Refer to 3.11, “Managing the Hardware Resource Configuration” on page 53, for more details on how to do this.

14. Print the system configuration and store it with the service and disaster recovery procedures.

    Refer to 3.11.4, “Printing the System Configuration” on page 58, for details on how to do this.

2.4 Hardware Considerations

Every logical partition requires mandatory hardware to function properly. Logical partitions can also be configured with other optional hardware, some of which can be switched between partitions. Before planning the detailed hardware requirements for a system, it is important to understand the basic hardware requirements.

Partitioning can only be implemented on an Nway AS/400 processor. Each partition requires:

- A minimum of one processor
- A minimum of 64 MB of memory (256 MB in the primary partition)
- A console device for each partition
- A load source DASD
- An alternate IPL device (this may be a device that can be shared with other partitions)
- IOPs that support the above three devices

Full details of the hardware supported with partitioning can be found in the *Logical Partitions Hardware Planning Guide* accessible from the logical partitioning home page at: `http://www.as400.ibm.com/lpar`

2.4.1 Supported AS/400e Models

Partitioning is only supported on the AS/400e series hardware. This restriction is due to processor architecture on earlier AS/400 models.

Table 2. Supported Models

AS/400 Model Number	Processor Feature Codes
AS/400e server S20	#2165, #2166
AS/400e Custom Mixed-Mode server S20	#2170, #2177, #2178
AS/400e server S30	#2258, #2259, #2260
AS/400e Custom Mixed-Mode server S30	#2320, #2321, #2322
AS/400e server S40	#2256, #2261, #2207, #2208
AS/400e Custom Mixed-Mode server S40	#2340, #2341
AS/400e model 620	#2182
AS/400e model 640	#2238, #2239
AS/400e model 650	#2240, #2243, #2188, #2189

AS/400 Model Number	Processor Feature Codes
AS/400e server 720	#2063, #2064
AS/400e server 730	#2066, #2067, #2068
AS/400e server 740	#2069, #2070

Partitioning offers IOP-level partitioning, as well as bus-level partitioning. The IOPs used on a shared BUS must be able to IPL independently. For a current list of supported IOPs, refer to the *Logical Partitions Hardware Planning Guide* (available from the logical partition home page).

During hardware planning, and with the movement and placement of IOP, you may need to purchase additional storage expansion towers or system expansion towers. Each additional tower is considered a separate bus. Ensure that the total number of buses across the entire system is not exceeded.

2.5 Software Considerations

Logical partitioning is supported in Version 4 Release 4 Modification 0 (V4R4M0) and subsequent versions of the OS/400 operating system. It is not supported on earlier releases of OS/400. Any systems that implement logical partitions in this first stage will run all partitions at V4R4M0 of OS/400.

2.5.1 Licensed Program Products — Licensing

A customer requires only one license of OS/400 per physical system, regardless of the number of partitions created. For licensed program products (LPP) that are licensed by processor group, the customer purchases one license according to the processor group to which the physical system belongs, regardless of the number of partitions created. The LPP can then be installed in each partition on the system if required.

For LPPs that are licensed based on the number of users, the customer purchases a license according to the number of users required across the entire AS/400 system. The LPP can then be installed in each partition on the system if required.

In V4R4M0, license management for user-based LPP is checked only at the individual partition level. In future releases, this will be done at the system level, meaning the total of all users of the LPP in all partitions.

The following LPPs have user-based pricing:

- 5769-CL3 VisualAge RPG and Cooperative Development Environment for AS/400
- 5769-CP4 Call Path Server for AS/400
- 5716-CX4 VisualAge C++ OS/2 Client
- 5769-CX5 (Option 1) VisualAge C++ Windows 95/NT Client
- 5769-DFH CICS for AS/400
- 5769-FW1 Firewall for AS/400
- 5769-RD1 (Option 5) OnDemand Server Feature
- 5769-SM1 System Manager for AS/400
- 5769-SS1 (Option 25) OS/400 NetWare Enhanced Integration
- 5769-SVD Netfinity AS/400 Manager for OS/2
- 5769-SVE Netfinity AS/400 Manager for Windows 95
- 5769-SV3 ADSTAR Distributed Storage Manager for AS/400
- 5798-TBT OfficeVision JustMail
- 5769-VI1 ImagePlus VisualInfo for AS/400
- 5769-VI1 (Option 2) ImagePlus VisualInfo Object Server
- 5769-WP1 OfficeVision for AS/400
- 5769-XW1 AS/400 Client Access Family For Windows
- 5769-XY1 AS/400 Client Access Family
- 5769-XZ1 OS/2 Warp Server for AS/400

Third-party products are licensed according to the policy of the individual vendors.

2.5.2 Software Problems That Can Affect All Partitions

Logical partitioning documentation may mention some software failures affecting the entire machine or other partitions. This refers to a software failure that is occurring in the Partition Licensed Internal Code (PLIC). The frequency of this happening is regarded as very low due to the thorough testing of all releases.

2.5.3 Installing V4R4M0

We recommend that V4R4M0 is installed and allowed to stabilize in the customer's environment before logical partitions are implemented. As with any upgrade to OS/400, the customer should run for at least one business cycle. This will ensure that, in the unlikely event any problems are encountered, they can be resolved before any further changes are made to the system.

2.5.4 Application Dependencies between Partitions

There are no application dependencies when using partitions. Each partition can operate totally independently of the others. If required, they can have inter-partition communication between selected partitions. There is nothing in the partitioning function that changes the way that application code runs on the system.

2.5.5 National Language Considerations

There are no special national language considerations for logical partitioning. Secondary languages are installed in the same way as they are today.

The situation may arise where a customer requires a different primary language for each partition created on the system. This is allowed under the terms of the licensing agreement. As the time this redbook was published, it was not known how this will be supported by customer fulfillment. For further information, please contact your local IBM representative.

Some countries choose to translate the dedicated service tool (DST) displays into the relevant national language. This may influence the primary language that is installed on a partition. The majority of logical partitioning configuration and maintenance tasks are performed from the primary partition dedicated service tool displays. However, some service-related tasks can only be performed from dedicated service tools on the secondary partitions.

2.5.6 Future Release Co-Existence

When planning for future release co-existence on the same physical hardware, consider the minimum level of OS/400 required to support any new hardware that becomes available.

2.5.7 Published Software Limits

The OS/400 software maximums are the limits per partition. For example, each partition can support up to a maximum of 16 auxiliary storage pools (ASPs).

2.6 Interface to Create and Manage Partitions

You can only restart, reconfigure, and manage logical partitions from the DST and system service tools (SST) environment. For information on which options can be accessed from each environment, please refer to Appendix C, “Accessing LPAR Functions” on page 175.

DSTs are a set of tools normally used to work with the system when the operating system is not available. SST are a subset of the DST tools. The tools available through SST, such as displaying the system configuration, can be used while the operating system is running and other users are on the system.

These environments do not provide the same user interface and support that is provided with standard OS/400 displays. For example, there is no cursor-sensitive help, second-level information available for messages, or prompting of options or commands. There are many system maintenance functions that can be performed from these menus that assume the user is aware of the impact that their actions will cause.

People who are not familiar with using this interface find the management of a partitioned environment very different from what they are used to using. The single, most important point to remember when using the DST or SST interface is the message that is presented when you first select the "Work with system partitions" option from the main DST or SST menu.

Attention

Incorrect use of this utility can cause damage to data in this system. See the service documentation that comes with the system. For more information on how to use DST and SST, access the IBM AS/400 Technical Studio at: `http://www.as400.ibm.com/tstudio`

If you are not familiar with the DST and SST environment, we recommend that you *do not attempt* to configure LPAR on your own. Assistance from IBM is available under the terms and conditions of a normal services contract.

The DST and SST interface is different from what the majority of customers use. We recommend that you take this into account when planning the level of IOP switching that is going to occur between partitions.

The parameter settings specified when setting up and creating partitions can have a major impact on the management and performance of the partitions. Refer to Chapter 3, “Operational Considerations” on page 39, to understand the full implications of these parameters.

2.7 Deciding What Runs in Primary and Secondary Partitions

As defined in 1.5, “Primary Partition” on page 12, a primary partition exists on every N-way AS/400e model. It is possible to create multiple secondary partitions depending on the resources available.

The primary partition is a fully functional system with additional responsibility to manage all secondary partitions. All partition management functions are included in the Partition Licensed Internal Code (PLIC) of the primary partition.

Secondary partitions are logical systems created from the hardware resources that are available to them. Secondary partitions are independent of each other, but maintain a dependency on the primary.

Due to the significance of the primary partition, take care when deciding what workload to run in that partition. You have three choices of the type of work that can run in the primary partition:

- **Nothing** — The primary partition is configured with the minimum amount of resources allowable and purely performs partition management tasks.
- **Production** — The primary partition is configured to run a production environment. The partition resources are allocated to allow for the system to run the selected workload, plus partition management tasks.
- **Test** — The primary partition is configured to run a test environment. The partition resources are allocated to allow for the system to run the selected test workload, plus the partition management tasks. This *is not* recommended due to the instability that normally exists in this kind of environment.

There is much discussion that can be had about the pros and cons of each choice that cannot be adequately covered here. The decision of what workload is run in each partition should take into account these aspects:

- **Stability of the application** — An application that has not reached stability or is going through frequent change can impact the partition in which it runs. Performance can be erratic if programs go into loops, or if errors are constantly being encountered that have to be handled by the operating system.
- **Availability requirements** — Customers who have a high availability requirement on any of the secondary partitions should not run applications in the primary partition. High availability planning requires that all possible steps are taken to minimize the impact any problems encountered can have on the applications requiring high availability. An application problem

encountered in the primary partition can impact the partitions performance, operating system, or LIC. Such a problem can also cause a system failure. A failure in the primary partition can, therefore, affect the secondary partition and the applications that are running in the secondary partition.

- **Application testing** — An error in an application function does not normally result in the failure of an AS/400 system. Errors encountered normally result in OS/400 exception messages being issued and the function ending abnormally. If the problem is found to be in IBM generated code, then a PTF is required to fix the problem. Testing applications that use leading edge functions that are likely to require frequent PTFs should not be done in the primary partition. The disruptive nature of installing PTFs can affect all secondary partitions. This is particularly true when PTFs require a system IPL on the primary partition.
- **IBM code testing** — New release testing or PTF testing *should not be done* in the primary partition if production applications exist in the secondary partitions. In very rare circumstances, defective PTFs can effect all secondary partitions. Using the primary partition for this sort of testing is similar to installing new releases of OS/400 or cumulative PTF packages directly onto a production machine.
- **Allocation of resources** — Allocation of resources can be performed from within the DST or SST environment. Incorrect use of the utilities on these menus can cause damage to data on the system. Consider the skill level of the personnel likely to require access to the DST and SST environment. Applications should not be run in the primary partition if it requires personnel who are unfamiliar with the DST or SST environment to perform maintenance through these screens.
- **Security of partitions** — You can restart, reconfigure, and manage all logical partitions from the primary partition by using DST or SST. If applications require personnel to have access to the DST or SST function, remember that they can affect all logical partitions. We strongly recommend that you change and secure DST and SST passwords.

2.8 Choosing Bus- or IOP-Level Partitioning

Depending on the requirements, there can be advantages to setting up one type of logical partitioning over another. There are two types of partitions that can be created:

- Bus-level partitions
- IOP-level partitions

For a full definition of each type, refer to the article *Planning for and Setting Up Logical Partitions* article. You can find it on the logical partition home page at: `http://www.as400.ibm.com/lpar`

Consider the points described in the following sections when deciding whether to have Bus or IOP level partitions.

2.8.1 Bus-Level Partitions

Bus-level partitions offer the following benefits:

- Better problem isolation between partitions since each bus and all the devices attached to that bus do not depend on other partitions
- Better performance since SLIC does not have to call PLIC to queue commands for a bus
- Less day-to-day management of IOP resources since they are permanently allocated to the partition
- Limited requirements for personnel unfamiliar with DST/SST environment to access DST functions

Note: The *total* number of buses that an AS/400 system can have is 19, regardless of how partitions are defined. You should plan accordingly.

2.8.2 IOP-Level Partitions

IOP-level partitions offer the following advantages:

- Allows resources to be switched among multiple partitions. This can help reduce the cost of hardware required.
- IOP resources need to be managed from the SST or DST environment, but do not require a system or partition IPL to become active. Incorrect use of the utilities on the DST and SST menus can cause damage to data on the system. Consideration should be given to the skill level of the personnel likely to require access to the DST and SST environment.
- Any bus problems encountered can potentially affect multiple partitions.

Note

Regardless of the type of partioning selected, the primary partition controls the availability of each bus on the system. If a secondary partition owns a shared bus, and that partition is powered off, all IOPs on that shared bus dedicated to other partitions can still be used.

2.9 Moving to a Partitioned Environment

A customer wanting to move from an existing environment to a partitioned environment needs to plan to minimize the impact to the existing users. You can do this by choosing one of the following options:

- **Install a new system.**

  The simplest approach is to install a new AS/400 system and set up the partitions ahead of time. The workload that is going to run in each partition can then be migrated at a convenient time to the new system. As each one of the existing systems is moved to the new system, the hardware from the existing system can be added to the new system ready for other partitions to use.

- **Replace the release on the current system.**

  The existing system is upgraded to V4R4M0 of OS/400 and the existing workload runs in the primary partition. After the stability of V4R4M0 is proven, secondary partitions can be created and workloads migrated to these new partitions. Careful planning is required since you may have to remove configured disks so that they can be assigned to the secondary partition.

- **Upgrade the current system.**

  The existing system is upgraded to V4R4M0 of OS/400 and the existing workload runs in the primary partition. After the stability of V4R4M0 is proven, additional hardware is added to the system. This hardware is used when creating secondary partitions. The workload can then be migrated to the new partitions.

2.10 System Naming Convention

The default system naming convention works the same way as it does today. The primary partition has a system name of Sxxxxxxx, where xxxxxxx is the seven-character system serial number that can be displayed using the following command:

```
DSPSYSVAL QSRLNBR
```

Since each system name generated must uniquely identify each partition within the physical system, all secondary partitions created will generate default system names in the form of Axxxxxxx through Kxxxxxxx.

The system name can be changed using the Change Network Attribute (CHGNETA) command. It does not depend on the partition name specified

when creating or changing a partition. However, to simplify matters, we recommend that you keep the partition name the same as the system name. A change to the partition name must be done from DST on the primary partition, but does not require an IPL.

2.11 System Serial Number

The system serial number is the same across all partitions. This is commonly used as a unique system identifier. The system serial number has to be qualified by the partition ID to achieve the equivalent effect.

The primary partition always has a partition ID of zero. Secondary partitions then start from 01 and increment according to the number that are created.

In Appendix B, “Sample RPG and C Programs” on page 171, there are two sample programs that show how to retrieve the individual partition information.

2.12 Security

You can only restart, reconfigure, and manage logical partitions from the DST and SST environment.

Systems that are set up to run partitions may require more access to the DST and SST environments than is typically necessary. As explained earlier in this chapter, these environments provide powerful tools that can cause major damage to your data if used incorrectly. We recommend that the default passwords for the DST and SST environment be changed to prevent unauthorized access.

We also recommend that the key is not left in the operators control panel. Otherwise, this would allow any unauthorized person to force the DST signon screen to the console of the primary partition.

Using non-partition related functions in DST and SST is restricted purely to the resources allocated to that partition. For example, it may not be possible for one partition to display or alter the contents of another partition’s memory or disk.

2.13 Updating Your Save Strategy

As you partition a system, you have to consider the requirements of all partitions when you develop your save and restore strategy. You should consider the following points when developing your strategy.

Consider Using a Tape Library
Having a physical system with multiple partitions and possibly switchable tape resources between these partitions is a more complex environment to manage from the save and restore perspective. We recommend that you consider a tape library and BRMS/400 as a way to help manage the new environment.

Availability of the Tape Drives
Different applications have different save and restore requirements. Because a system has a certain number of tape drives attached to it, does not mean that all tape drives are available for all partitions to use at anytime. If you switch removable media devices between partitions, then your backup of these partitions will become a sequential process.

Accessing More Hardware
It may be possible to have more hardware available for you to use on your save or restores.

The top half of Figure 12 on page 34 shows that System A, System B, and System C all have connections to the four 3590 model B11 tape devices. You can connect a tape unit 3590 Model B11 to a maximum of two AS/400 systems using the dual port on the tape control unit. The only system that can access all four tape drives is system B.

If these three systems were consolidated onto a single AS/400 system but with three partitions, as shown in the lower half of Figure 12, it would be possible for all tape drives to be available to each partition. Then, the tape drives have to be connected to the system using shared buses so that all partitions could use them.

Continue to follow any configuration recommendations for performance to ensure the same performance is achieved on a partitioned system as it would be on a traditional system. For further information, refer to the *AS/400 Performance Capabilities Reference Guide.* You can request a copy of this document by downloading it from the AS/400 On Line Library through the AS/400 Internet site at: `http://www.as400.ibm.com`

The document is viewable and downloadable in the Adobe Acrobat (.pdf) format and is approximately 1.5MB in size. The Adobe Acrobate Reader plug-in is available at: `http://www.adobe.com`

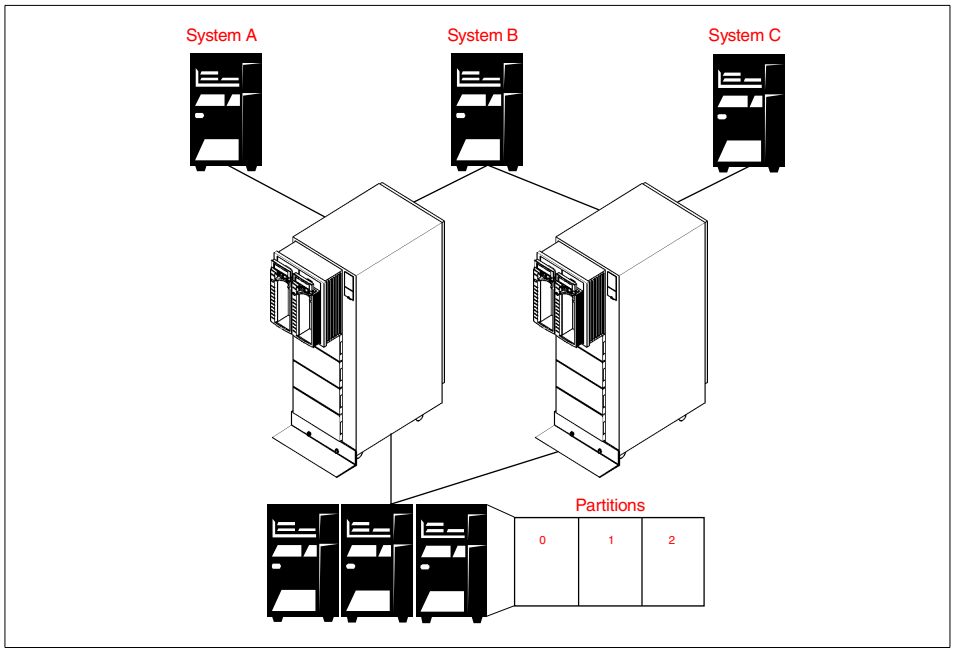


Figure 12. Using Multiple Tape Drives

Taking Saves from a Second Copy of the Data

Using one of the High Availability Business Partner (HABP) applications, it is possible to replicate data from one partition to another on the same machine. Although this option does not offer any higher level of availability with unplanned outages, it offers a higher level of availability with planned outages such as daily backup. The production machine can always be available to the users, while the second copy of data is used for the backup to tape.

The CPU required for this partition will have to be adequate so that data changes from other partitions can be received and applied and when required, the save tasks can be run. A save task is typically I/O intensive so this in itself will require very little CPU cycles.

Enough DASD needs to be installed in this partition to support the amount of data it will be required to store. This is likely to be the largest partition on the system.

When a save is required, the apply process in the backup partition is stopped and the save taken from this copy of the data. Work continues to run in the original partition and journal entries generated by changes to the data are still sent to the backup partition but not applied.

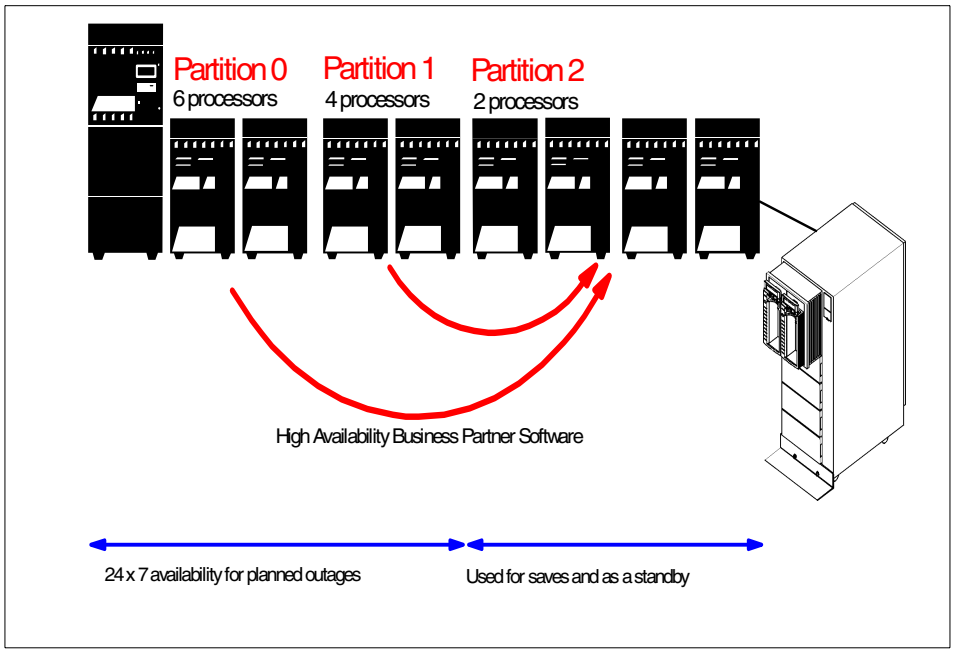


Figure 13. Saving from a Second Copy of the Data

Using ObjectConnect

ObjectConnect/400 saves information from one system and restores it to another using a single command. ObjectConnect is a non chargeable product option of OS/400. For more information on ObjectConnect/400, refer to 7.2.2, “ObjectConnect/400” on page 137.

ObjectConnect can be used to snapshot data at a particular point in time. The users stop using the application so a copy of the data can be transferred to another partition. Later, this data can be saved to removable media.

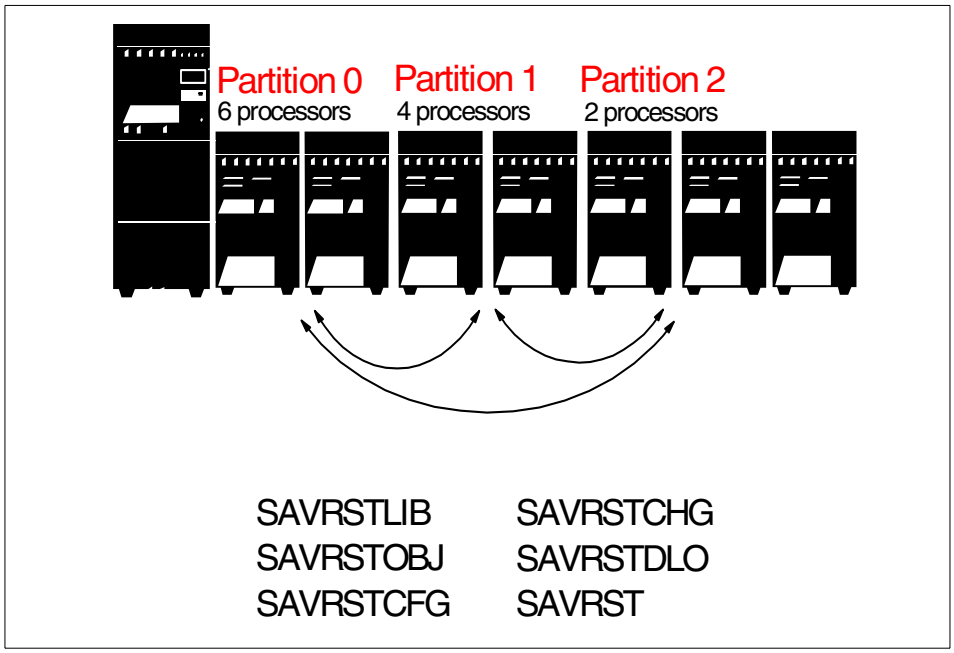


Figure 14. Using ObjectConnect to Move Data to Another Partition

The following test results demonstrate the gigabyte-per-hour (GB/HR) transfer rates that can be achieved when transferring data between partitions using ObjectConnect/400 and the virtual bus adapter. OptiConnect/400 or the OptiMover PRPQ (5799-FWQ) must also be installed to enable ObjectConnect to use the virtual high-speed bus adapter. Inter partition communication must be enabled on both partitions before this can be used. For more information, refer to Chapter 7, “Inter-Partition Communication” on page 133.

The tests were performed on a model 720 #2064 divided into two partitions. Each partition had two processors and 4096 MB of main storage.

Test 1 used one SAVRSTLIB command to save a single library from one partition and restore it to another. The library contained 22 objects and was 2.7 GB in size. This job transferred the data at the equivalent rate of approximately 27 GB/HR.

Test 2 used one SAVRSTLIB command to save 20 libraries from one partition and restore them to another. Each library contained 22 objects and the total amount of data transferred was 54 GB. This job transferred the data at the equivalent rate of approximately 28 GB/HR.

Test 3 used twenty concurrent SAVRSTLIB commands to each save a single library from one partition and restore it to another. Each library contained 22 objects. The total amount of data transferred was 54 GB. These jobs transferred the data at the equivalent rate of approximately 28 GB/HR

As the results show, it is possible to consistently achieve a save and restore rate similar to what a 3590 tape drive can achieve.

2.14 Published Hardware Limits

The AS/400e hardware maximums published in the *AS/400e System Handbook*, GA19-5486, are the limits per physical system. The sum of individual hardware types across all partitions must not exceed the published limit for a physical system.

2.15 DASD Compression Support

DASD compression support is the same on a partitioned system as it is on a non-partitioned system. Compressed DASD is only supported in user ASPs. DASD compression must be turned off for a disk or disk IOP *before* it is removed from a user ASP, if the DASD is going to be used for the system ASP on another partition.

2.16 Dual Tape Attach Support to Multiple Partitions

It is possible to attach some model tape devices to two AS/400 systems. Only one system at a time can use the device.

On a partitioned system, it is possible to attach these drives to two partitions using separate IOPs. In the event that these IOPs are subsequently dedicated to the same partition, the same usage rules apply.

The system only allows one device description to be varied on at any one time. This allows the second device description that uses the second IOP to remain inactive.

Chapter 3. Operational Considerations

This chapter offers useful information to help you get ready for and run a system with multiple partitions.

3.1 Installing New Hardware Ready for Partitioning

The Technology Solutions Center (TSC) provides a list of any new hardware that is required to support the planned partitioning environment. This hardware needs to be installed before the partitions can be created.

The IOP placement information is also provided by the TSC. The hardware placement information should be reviewed with the customer engineer (CE) prior to installing the new hardware. This ensures that it is installed in the correct place according to the specific partition requirements.

The following sections explain how the hardware is treated by the system when it is installed. This assumes V4R4M0 has been installed.

Processors and Main Storage
If processors or main storage are added to a system that has no other partitions created, other than the default of the primary, the new hardware will be added to the primary partition. If processors or main storage are added to a system with secondary partitions, the new hardware needs to be allocated to a partition before it can be used.

Interactive Performance
If the customer adds interactive performance (model 7xx) to a system, the interactive performance is available immediately. Interactive performance is assigned by allocating a percentage of the total to each partition. For a system with no partitions created, 100% is allocated to the default primary partition.

The percentage amount allocated to each partition should be recalculated after the upgrade. This ensures that the amount allocated is still correct in light of the larger interactive performance installed.

Bus
Buses added to the system are added to the primary partition with a status of "own bus shared". For more information on this term, refer to A.13, “Ownership and Sharing” on page 169.

Important

Before installing the hardware, it is important to review the output from the capacity plan. Any specific recommendations can be implemented during the allocation of hardware resources when creating the partitions.

IOP

If IOPs are added to a system with multiple partitions defined. The new hardware will be dedicated to the partition that owns the BUS.

3.2 Temporarily Making New Hardware Unavailable

If the new hardware is to be added to an existing system, it is possible that the customer will not want the new hardware to be used until the partitioned environment is established. Depending on the type of resource, the available options are different.

3.2.1 Processors

To make processors temporarily unavailable, use the techniques described here:

On the Change Partition Processing Resources display, it is possible to change the number of processors available to a partition. If after installing the new processors, the number available to the primary partition is reduced. This places the additional processors into a ready, but not allocated, state. When the secondary partitions are created later, these processors can then be assigned to that partition.

The advantage of this approach is to ensure that correct performance expectations are set with the users. It also ensures that they do not see a performance penalty when the processor or any other resource is removed from the partition in which they are working.

```
                Change Partition Processing Resources
                                                        System:   SYSTEMA
Type changes, press Enter.

  Partition identifier . . . . . . . . . . . . . . :  1
  Partition name . . . . . . . . . . . . . . . . . :  PRIMARY

  Current / available number of processors . . . :  2    / 2
  New number of processors . . . . . . . . . . . .  2
  Minimum / maximum number of processors . . . . .  1    / 2

  Current / available size of main storage (MB)  :  4096     / 4096
  New size of main storage (MB)  . . . . . . . . .  4096
  Minimum / maximum size of main storage (MB)  . .  2048     / 8192

  Current / available interactive performance  . :  46  / 4   %
  New interactive performance  . . . . . . . . . .  46  %
  Minimum / maximum interactive performance  . . .  25  / 46  %

  Connect to system inter-partition OptiConnect  .  1  1=Yes, 2=No

F3=Exit                                          F9=Exclude limits
F11=Display partition processing configuration   F12=Cancel
```

Figure 15. Change Partition Processing Resources — Processors

3.2.2 Main Storage

To make main storage temporarily unavailable, use one of the following techniques:

- **Option 1**

  On the Change Partition Processing Resources screen, you can change the amount of main storage available to a partition. If after installing the new main storage, the customer reduces the amount available to the primary partition, the additional main storage will be placed into a ready, but not allocated, state. When the secondary partitions are created later, the main storage can then be allocated to that partition.

```
                Change Partition Processing Resources
                                                        System:   SYSTEMA
Type changes, press Enter.

  Partition identifier . . . . . :  1
  Partition name . . . . . :  PRIMARY

  Current / available number of processors . . . :  2   / 2
  New number of processors . . . . . .  2
  Minimum / maximum number of processors . . . . .  1   / 2

  Current / available size of main storage (MB)  :  4096     / 4096
  New size of main storage (MB)  . . . . . .  4096
  Minimum / maximum size of main storage (MB)  . .  2048     / 8192

  Current / available interactive performance  . :  46  / 4   %
  New interactive performance  . . . . . .  46  %
  Minimum / maximum interactive performance  . . .  25  / 46  %

  Connect to system inter-partition OptiConnect  .  1  1=Yes, 2=No

F3=Exit                                          F9=Exclude limits
F11=Display partition processing configuration   F12=Cancel
```

Figure 16. Change Partition Processing Resources — Main Storage

- **Option 2**

  Create and start a new subsystem description that allocates the extra main storage when the subsystem is started. No workstation or job queue entries should be specified to prevent the system from routing any work through the subsystem.

3.2.3 Interactive Performance

To make interactive performance temporarily unavailable, use the following technique:

On the Change Partition Processing Resources display, you can change the amount of interactive performance available to a partition. If after installing the new interactive performance the percentage available to the primary partition is reduced, the additional performance is prevented from being used by the primary partition. When the secondary partitions are created later, the available interactive performance can be assigned to these partitions.

```
                Change Partition Processing Resources
                                                        System:   SYSTEMA
Type changes, press Enter.

  Partition identifier . . . . . :  1
  Partition name . . . . . :  LPAR2

  Current / available number of processors . . . :  2   / 0
  New number of processors . . . . . .  2
  Minimum / maximum number of processors . . . . .  1   / 2

  Current / available size of main storage (MB)  :  4096     / 0
  New size of main storage (MB)  . . . . . .  4096
  Minimum / maximum size of main storage (MB)  . .  2048     / 4096

  Current / available interactive performance  . :  46   / 4    %
  New interactive performance  . . . . . .  46   %
  Minimum / maximum interactive performance  . . .  25   / 46   %

  Connect to system inter-partition OptiConnect  .  1  1=Yes, 2=No

F3=Exit                                          F9=Exclude limits
F11=Display partition processing configuration   F12=Cancel
```

Figure 17. Change Partition Processing Resources — Interactive Performance

3.2.4 I/O Resources

To make I/O resources unavailable, use one of the following techniques:

- **Option 1**

  Remove the I/O resource from the owning partition's configuration. This makes them available to be allocated when the secondary partitions are created.

- **Option 2**

  Do not create or vary off the configuration descriptions that use these resources until they are required.

3.3 Moving Existing IOPs

If partitioning is going to be implemented on an installed system, it may be necessary to physically move some IOPs to support the new partitioned environment. Take care when moving IOPs to a different position on the same system. In certain circumstances, new resource descriptions are created by the AS/400 system. In this case, any configuration objects that are created need to be changed to refer to the new resource description.

3.4 Retrieving the Resources Allocated to a Partition

In V4R4M0, there is no API to access partition resource information. There is an unblocked Machine Instruction (MI) that can be used in a high-level language program to retrieve this information.

The MI returns the following information about a partition:

- The number of partitions on the system
- The partition number from where this information was retrieved
- The logical serial number for that partition
- The minimum number of processors configured
- The maximum number of processors configured
- The current number of processors configured
- The minimum amount of memory configured
- The maximum amount of memory configured
- The current amount of memory configured
- The minimum percentage of interactive performance configured
- The maximum percentage of interactive performance configured
- The current percentage of interactive performance configured

Refer to Appendix B, “Sample RPG and C Programs” on page 171, for sample RPG and C programs. These programs show how to use this unblocked MI instruction.

3.5 Changing Partition Processing Resources

Once a partition is created, there are some changes to the configuration that, if made, require an IPL of the affected secondary partition or an IPL or the primary (the entire system) for the changes to come into effect. If changes are made to the number of processors allocated, the amount of memory allocated or the amount of interactive performance allocated to a partition and the new value is still within the minimum or maximum values for that partition, only the affected partition needs to be IPLed. If the new value allocated requires the minimum or maximum values to be changed, this requires an IPL of the primary partition before it comes into effect.

If any partition has a changed "Connect to System Inter-Partition OptiConnect" parameter, then an IPL of the primary partition is required before this change becomes activated.

If a partition requires an IPL to put into effect some changes that were made to the processing resources, a "<" symbol is shown next to that partition on

the Work with Partition Configuration screen. If these changes are reversed before an IPL, the symbol is removed from the display.

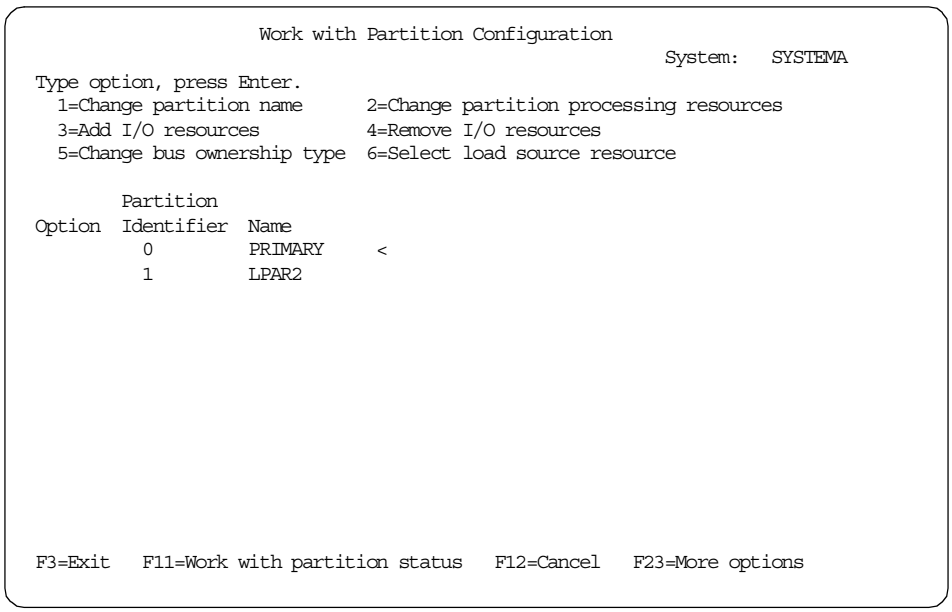


Figure 18. Work with Partition Configuration Display with Changes Pending

3.6 Saving and Restoring Partitions

Introducing partitions into a customer environment may add new steps into the customers save and restore process. It should also initiate a review of the entire save and restore strategy.

There is no way to perform a single save or restore for a physical system. Each partition needs to be treated separately, but it is possible to perform multiple saves or restores in different partitions at the same time, assuming they each have their own save and restore device. In principle, the process of saving or restoring logical partitions is the same as saving or restoring a system without logical partitioning.

You can use the Save Storage (SAVSTG) command to save partitions data. Remember that this command causes the system to power down and start the system again as if PWRDWNSYS RESTART(*YES) was specified. If you are going to use this command on the primary partition, make sure that all secondary partitions are powered off before you start.

3.6.1 Saving the Logical Partition Configuration

You cannot save the configuration data for logical partitions to a removable media device. The system automatically maintains the configuration data for all logical partitions created on a physical system on the load source disk of each partition.

When the system is IPLed, it validates the configuration available on the load source disk of the primary partition. If this is not available, it searches the other load source disks and prompts you to select the configuration that should be used.

On a scratch installation of the primary partition, the partition configuration must be recovered before the recovery of the disk configuration. For more information on the steps that are required, refer to *Backup and Recovery,* SC41-5304.

Not saving the partition configuration with the save commands allows data to be restored to a system whether it has logical partitions. A full system save can be used to restore to:

- The same logical partition
- A different logical partition on the same system
- A different logical partition on a different system
- A system that does not have any partitions at all

The save and restore process uses the machine serial number to determine if an object is being restored to the system from which it was saved. Starting with V4R4M0, a save or restore uses the serial number and partition number to determine this. Objects that were saved prior to V4R4M0 and are restored to a V4R4M0 system assumes that they are being restored to the same system if the serial number matches.

3.6.2 Saving Data from Logical Partitions

All save operations operate on the logical partitions as if they were separate systems. Using the save commands available, a customer may save their logical partition as they do today. The save will contain only the information from that logical partition.

3.6.3 Restoring Data to Logical Partitions

The restore commands available to customers may restore their logical partition data the same way as they would restore data to their system today. The commands that are used place the data into the partition from where the restore command is issued.

3.7 System Values Affected by Logical Partitioning

Every partition operates independently of the others. Most system values are unique to that partition. However, there are some system values that can only be changed from the primary partition. The value to which they are set influences how the secondary partitions will behave in certain circumstances.

3.7.1 QPRCMLTTSK — Processor Multi-Tasking

The QPRCMLTTSK system value allows you to turn on and turn off the processor multi-tasking capability. If enabled, more than one set of task data will reside in each CPU. Some workloads may experience increased performance due to caching implications. This system value setting affects all processors in all partitions.

3.7.2 QPWRRSTIPL — Automatic IPL after Power Restored

The QPWRRSTIPL value specifies whether the system should automatically IPL when utility power is restored after a power failure. It only affects the primary partition. Whether a secondary partition is IPLed at the same time as the primary partition depends on the secondary partition's configuration value for System IPL Action.

If the secondary partition System IPL Action is set to IPL, the partition automatically performs an IPL when the primary partition is IPLed. This is set from the Work with Partition Status display on the primary partition.

```
                        Work with Partition Status
                                                              System:   SYSTEMA
Type options, press Enter.
 13=IPL partition on system IPL   14=Hold partition on system IPL
 20=Display system type/model     21=Force Dedicated Service Tools
 22=Force Main Storage Dump       34=Force CPM or MSD IPL retry

     Partition             IPL     IPL                  Sys IPL  Reference
Opt  Identifier  Name      Source  Mode     State       Action   Codes
 __    0         PRIMARY   B       Normal   On          IPL      11
 __    1         LPAR2     B       Manual   On          IPL      11
 __    2         LPAR3     B       Manual   Off         HOLD     11









F3=Exit   F5=Refresh                    F10=Monitor partition status
F11=Work with partition configuration   F12=Cancel   F23=More options
```

Figure 19. Work with Partition Status Display

3.7.3 QRMTIPL — Remote Power On and IPL

The QRMTIPL value specifies whether remote power on and IPL can be started over a telephone line. It only affects the primary partition. Whether a secondary partition is IPLed at the same time as the primary partition depends on the secondary partition's configuration value for System IPL Action.

If the secondary partition System IPL Action is set to IPL, the partition automatically performs an IPL when the primary partition is IPLed. See Figure 19 for details on setting this parameter.

3.7.4 QUPSDLYTIM — Uninterruptible Power Supply Delay Time

The QUPSDLYTIM value specifies the amount of time that elapses before the system automatically powers down following a power failure. When a change in power activates the Uninterruptible Power Supply (UPS), messages are sent to the UPS message queue (the system value QUPSMSGQ). This system value is only meaningful if your system has a battery power unit or has an uninterruptible power supply.

Once the primary partition receives notification of a power failure, it notifies the secondary partition that such a condition exists. The messages are sent to the UPS message queue specified in the QUPSMSGQ on the secondary partitions. The value specified in the primary partition is the value that the system will use for all partitions.

It is important that you monitor for power failures across all partitions, and not just primary partitions. This ensures that the secondary partitions (which may be running a production workload) perform a normal shutdown in the event of a power failure.

Note

If any of the previously described system values are displayed or retrieved from a secondary partition, the value always reflects what was specified in the primary partition.

3.7.5 Other Related System Values and Commands

The following list of system values and commands are not enforced by the operating system. However, their values can impact operating partitions.

- **QAUDENDACN — Auditing end action**

  Contains the action that should be taken by the system when audit records cannot be sent to the auditing journal because of errors that occur when the journal entry is sent. If this is set to *PWRDWNSYS, the system ends should the attempt to send the audit data to the security audit journal fails. It ends with a B900 3D10 system reference code.

  If this system value is set to *PWRDWNSYS in the primary partition, then this would effect all secondary partitions should a problem be encountered.

- **QCMNARB — Communication arbiters**

  Contains the number of communication arbiter system jobs that are available to process work for controllers and devices. If set to *CALC (the shipped value), the operating system calculates the number of communication arbiter system jobs to be equal to the number of processors on the system.

  On a partitioned system, the number of jobs started are equal to the number of processors on the physical system, not the number assigned to that partition. We recommend that this system value be set manually.

- **QIPLDATTIM — Date and time to automatically IPL**

  Specifies a date and time when an automatic IPL should occur. This system value can be set independently in each partition. If the primary partition is powered down at the time that an automatic IPL should occur in a secondary partition, the IPL will not occur. When the primary partition performs an IPL, the secondary partition is IPLed if its IPL date and time are past due. The secondary partition performs an IPL even if it was configured with a System IPL Action of Hold.

  The power on and off schedule that can be set from the POWER menu also uses the system value QIPLDATTIM. This function works the same way as described in the previous paragraph.

- **QMODEL — System model number**

  This is the number or letter used to identify the model of the system. You cannot change QMODEL, but the four-character value can be displayed or retrieved in user-written programs. The system model number is the same in each partition on a system.

- **QPRCFEAT — Processor feature**

  This is the processor feature code level of the system. You cannot change QPRCFEAT, but the four-character value can be displayed or retrieved in user-written programs. The processor feature system value is the same in each partition on a system.

- **QSRLNBR — System serial number**

  This value cannot be changed. It is retrieved from the data fields by the system when installing the OS/400 licensed program. You can display QSRLNBR, or you can retrieve this value in user-written programs. The system serial number is the same in each partition on a system.

- **CHGIPLA Command**

  The Change IPL Attributes (CHGIPLA) command allows you to change the settings of attributes that are used during the initial program load (IPL).

  The "Hardware diagnostics (HDWDIAG)" attribute on this command specifies whether certain hardware diagnostics should be performed during the IPL. The list of diagnostics is pre-determined by the system and cannot be modified by the user.

  Hardware diagnostics can only occur when the primary partition is IPLed. This attribute cannot be changed from a secondary partition. Attempting to change this attribute in a secondary partition will result in message CPF18B4 being issued.

3.8 Security Considerations

It is possible to have different OS/400 security levels (QSECURITY) in each partition and a different security policy implemented by the customer for each partition. One partition may have security down to the object level, but another may have all objects secured by group profiles.

A user profile is required for each partition on a physical machine. System-supplied user profiles will exist on each partition after the operating system is installed. However, there is no dependency between partitions for these profiles.

It is possible for user profiles, for example QSRV, to have different passwords for each partition on the physical system. Typically, most maintenance requiring access by an IBM service representative is done from the primary partition. All QSRV passwords should be made available in case they are required.

When configuring partitions, it is possible to specify that *Interpartition OptiConnect* should be enabled. Any changes made to this parameter after the initial creation requires a full system IPL. For details on how to restrict user authority to OptiConnect, refer to the manual *OptiConnect for OS/400,* SC41-5414.

3.9 Date and Time Processing between Partitions

Each partition has its own system date and time as set by the QDATE and QTIME system values. There is only one clock on each physical system. This stores the value of QDATE and QTIME from the primary partition. To ensure that the secondary partitions date and time are stored correctly, each secondary partition maintains a delta value to the system clock. They use this to set the partition date and time when the partition is IPLed.

If the date or time is changed in the primary partition, the service processor updates the delta in each secondary partition to maintain the correct date and time for the secondary partitions. Each time that a change is made to the date or time in a secondary partition, the virtual service processor updates the delta accordingly.

If a hardware problem occurs in the primary partition that requires the date and time to be entered on the next IPL, the system automatically signals all secondary partitions that their date and time values are invalid. On the IPL of

the secondary partitions, the correct date and time for each partition must be entered when prompted to do so.

It is possible to display the reference code history of the secondary partitions from the primary partition. The date and time displayed indicate when the primary partition detected the reference code. They are not the date and time from the secondary partition settings.

```
            Display Secondary Partition Reference Code History
                                                          System:   SYSTEMA
Secondary partition(s) to display . . . . .  *ALL  *ALL, 1-1
Number of reference codes to display  . . .  200   1-200

Partition
Identifier  Name      Reference Codes  Date      Time
  1         LPAR2     11              01/24/99  20:56:45
                      11 C900 2F00    01/24/99  20:56:45
                      11 C900 2C25    01/24/99  20:56:44
                      11 C900 2C20    01/24/99  20:56:44
                      11 C900 2C40    01/24/99  20:56:44
                      11 C900 2C10    01/24/99  20:56:13
                      11 C900 2B40    01/24/99  20:56:12
                      11 C900 2B30    01/24/99  20:56:11
                      11 C900 2B10    01/24/99  20:56:02
                      11 C900 2AC0    01/24/99  20:56:01
                      11 C900 2AB0    01/24/99  20:56:01
                      11 C900 2AA5    01/24/99  20:56:00
                      11 C900 2AA4    01/24/99  20:56:00
                                                          More...
F3=Exit   F5=Refresh              F6=Print
F9=Include reference code detail  F12=Cancel
```

Figure 20. Display Secondary Partition Reference Code History

3.10 Main Storage Dump Processing

The main storage for each partition is specified on the Create Partition or Change Partition Processing Resources display. The primary partition requires a minimum of 256 MB of main storage. Secondary partitions require a minimum of 64 MB. Creating a partition with only the minimum amount of main storage specified allows only a very limited workload to run in that partition.

Main storage can be allocated in 1 MB increments over and beyond this minimum. This can result in one physical main storage card being allocated to two partitions. This is not a problem.

For completeness purposes, the main storage dump (MSD) process is explained in the following steps to show how the system copes with this type of configuration:

1. Once a fail condition is detected, the secondary logical partition sends a signal to the virtual service processor.
2. In the primary partition, the signal to the virtual service processor also results in the SRC being written to the log.
3. All IOPs that are dedicated to the failing partition are reset.
4. The IOP that has the load source DASD attached to it is IPLed by the primary partition.
5. The secondary partition copies 128 MB of memory to the load source DASD and continues to IPL to DST with that 128 MB of memory.
6. The secondary partition performs MSD processing and displays the MSD screens.
7. Once MSD processing is complete, the partition is IPLed.

A failure in the primary partition will cause all secondary partitions to end abnormally. Extra options were added to the main storage dump process display to allow a specific main storage dump, or all of them, to be selected.

3.11 Managing the Hardware Resource Configuration

You can display the hardware resource configuration from the DST, SST, and logical partition displays. Where you view the configuration from determines the level of detail that is displayed. This section explains the information available from each display.

3.11.1 Hardware Resource Information from DST and SST

Use the Hardware Service Manager within DST or SST to display the installed hardware resources. The hardware service manager is typically used by service representatives when performing hardware maintenance on the system.

The processors, main storage, and bus adapters that are displayed are for the entire system, not just the partition from where the option is being viewed. The IOP resources displayed are those currently reporting to the partition from where the option is being viewed.

Failed and non-reporting resources can also be displayed. These resources, at one time, were reported as operational for that partition.

It is possible to print the configuration displayed by selecting F6 (Print configuration) from the Hardware Service Manager initial menu (see Figure 21).

```
                          Hardware Service Manager

Attention:  This utility is provided for service representative use only.

   System unit . . . . . :   9406-720 10-394TM
   Release . . . . . :   V4R4M0 (1)

Select one of the following:

   1. Packaging hardware resources (systems, frames, cards,...)
   2. Logical hardware resources (buses, IOPs, controllers,...)
   3. Locate resource by resource name
   4. Failed and non-reporting hardware resources
   5. System power control network (SPCN)
   6. Work with service action log
   7. Display label location work sheet
   8. Device Concurrent Maintenance

Selection  2


F3=Exit      F6=Print configuration      F9=Display card gap information
F10=Display resources requiring attention      F12=Cancel
```

Figure 21. Hardware Service Manager — Initial Menu

```
                Logical Hardware Resources on System Bus

System bus(es) to work with . . . . . *ALL  *ALL, 1-  2
Subset by . . . . . . . . . . . . . . . *ALL  *ALL, *STG, *WS, *CMN, *CRP

Type options, press Enter.
  2=Change detail    4=Remove    5=Display detail    6=I/O Debug
  8=Associated packaging resource(s)     9=Resources associated with IOP

                                                                  Resource
Opt  Description                    Type-Model  Status           Name
     System Bus                          -      Operational      SPD01
      Multiple Function IOP    * <   9164-001    Operational      CMB01
     System Bus                          -      Operational      SPD02
     Virtual System Bus                  -      Operational      SPD03
      Virtual Bus Adapter           268A-000    Operational      SOC01




F3=Exit    F5=Refresh    F6=Print    F8=Include non-reporting resources
F9=Failed resources      F10=Non-reporting resources
F11=Display serial/part numbers      F12=Cancel
```

Figure 22. Hardware Service Manager — Sample Resources Display

3.11.2 Hardware Resource Information from Logical Partitions

You can only access I/O resource information from within the logical partitioning displays. The amount of information that is available depends on whether you are viewing it from the primary or secondary partition. The I/O resource information for each partition is rebuilt when a partition IPLs.

If an IOP within a tower is replaced using the concurrent maintenance facility, the I/O resources are not updated properly until the partition is IPLed. If a partition is currently powered off, the I/O resources displayed reflect the configuration at the time the partition was powered off.

When switching I/O resources between partitions, a time delay may exist before the resource is displayed as allocated to the partition. This is a result of the time it takes the IOP to IPL and update the resource information.

3.11.2.1 Information Available from the Primary Partition

To display the allocated I/O resources for all partitions in partition ID sequence, perform the following steps from the Work with System Partitions initial menu:

1. Select option **1** (Display partition information).
2. Select option **3** (Display allocated I/O resources).

```
                    Display Allocated I/O Resources
                                                          System:   RCHASM25
System partition(s) to display . . . .  *ALL  *ALL, 0-1
Level of detail to display . . . . . .  *ALL  *ALL, *BUS, *IOP, *IOA, *DEV

Par                                              Serial      Part
ID   Description                    Type-Model   Number      Number
  0  PRIMARY  Partition             9406-720     10-394TM
     System Bus                                  00-0000000
      Multiple Function IOP   *<    9164-001     10-8285313  0000087G0725
       Communications IOA           2720-001     53-8302889  0000090H9189
        Communications Port         2720-001     53-8302889  0000090H9189
       Communications IOA           2850-011     10-8282050  0000061H0127
        Communications Port         2724-001     10-8215054  0000044H7571
        Communications Port         2838-001     10-8316090  0000021H5458
        Virtual Port                6B00-001     10-8282050  000008193654
       Communications IOA           285A-003     10-8301085  0000090H9212
       Workstation IOA              266C-001     53-8302889  0000090H9189
        Display Station             3487-0HC     00-*******
                                                                    More...
* Indicates load source.
F3=Exit   F5=Refresh   F6=Print   F10=Display logical address
F11=Display partition status      F12=Cancel
```

Figure 23. Display Allocated I/O Resources from Primary Partition

To display the available I/O resources that can be allocated to the primary partition, perform the following steps from the Work with System Partitions initial menu:

1. Select option **1** (Display partition information).
2. Select option **4** (Display available I/O resources).

The available resources are attached to the system and are not allocated to another partition. For example, you may have a tape drive that will be shared across all of the partitions.

To display the total system I/O resources, perform the following steps from the Work with System Partitions initial menu:

1. Select option **1** (Display partition information).
2. Select option **5** (Display system I/O resources).

```
                    Display System I/O Resources
                                                          System:   RCHASM25
Level of detail to display . . . . . *ALL  *ALL, *BUS, *IOP, *IOA, *DEV

I/O Resource                                Serial      Part
Description                  Type-Model     Number      Number
System Bus                                  00-0000000
 Multiple Function IOP   *<  9164-001       10-8285313  0000087G0725
  Communications IOA         2720-001       53-8302889  0000090H9189
   Communications Port       2720-001       53-8302889  0000090H9189
  Communications IOA         2850-011       10-8282050  0000061H0127
   Communications Port       2724-001       10-8215054  0000044H7571
   Communications Port       2838-001       10-8316090  0000021H5458
   Virtual Port              6B00-001       10-8282050  000008193654
  Communications IOA         285A-003       10-8301085  0000090H9212
  Workstation IOA            266C-001       53-8302889  0000090H9189
   Display Station           3487-0HC       00-*******
   Display Station           3487-0HC       00-*******
  Multiple Function IOA      2741-001       10-8308021  0000091H0271
   Disk Unit                 6713-074       68-0171DE9  59H6611
                                                                  More...
F3=Exit   F5=Refresh   F6=Print   F10=Display logical address
F11=Display partition status      F12=Cancel
```

Figure 24. Display System I/O Resources Display

It is possible to print the system I/O resource information by selecting the F6 key from the display.

3.11.2.2 Information Available from the Secondary Partition

If option 3 is selected from the Display Partition Information menu on the secondary partition, only the allocated I/O resources for the secondary partition are displayed.

If option 4 is selected from the Display Partition Information menu on the secondary partition, only the available I/O resources that can be allocated to the secondary partition are displayed.

3.11.3 Removing Hardware Resources from Logical Partitions

Whether you are creating logical partitions for the first time or switching some hardware resources between partitions, it is important to understand how the resources should be removed from the current partition. You want to ensure that the hardware resource table is updated correctly.

```
                          Remove I/O Resources
                                                              System:   RCHASM25
Level of detail to display . . . . . *ALL  *ALL, *BUS, *IOP, *IOA, *DEV

Partition identifier . . . . . :  0
Partition name . . . . . :  PRIMARY

Type options, press Enter.
  1=Remove  2=Remove and clear hardware resource(s)

     I/O Resource                                  Serial      Part
Opt  Description                    Type-Model    Number      Number
     System Bus                                   00-0000000
      Multiple Function IOP   *<    9164-001      10-8285313  0000087G0725
       Communications IOA           2720-001      53-8302889  0000090H9189
        Communications Port         2720-001      53-8302889  0000090H9189
       Communications IOA           2850-011      10-8282050  0000061H0127
        Communications Port         2724-001      10-8215054  0000044H7571
        Communications Port         2838-001      10-8316090  0000021H5458
        Virtual Port                6B00-001      10-8282050  000008193654
                                                                        More...
* Indicates load source.
F3=Exit   F10=Display logical address   F12=Cancel
```

Figure 25. Remove I/O Resources Display

If the allocation of the hardware resource is going to be permanently removed from that partition, remove it by using option 2. This ensures that the hardware resource tables are updated correctly.

If the hardware resource is going to be used in this partition again, then remove it by using option 1. This ensures that the hardware resource names previously assigned by the system are kept and reused later when the hardware resource is reallocated to the partition. While the hardware resource is not allocated to the partition, it will have a status of "Non Reporting" when viewed from the Hardware Service Manager.

It is possible to remove non-reporting hardware resources by using the Hardware Service Manager in DST and SST. Or, option 4 can be used from the logical partitioning Recover Configuration Data menu. This should only be done by experienced service personnel. Incorrect use of these options may result in valid resource names being incorrectly removed.

3.11.4 Printing the System Configuration

We recommend that, once logical partitions are initially setup and every time the configuration changes, you print the system configuration and keep it with the service and disaster recovery documentation.

Follow these steps to print the system configuration:

1. From the primary partition, start SST or DST.
2. From SST, select option **5** (Work with system partitions). From DST, select option **11** (Work with system partitions), and press **Enter**.
3. Select option **1** (Display partition information).
4. Select option **5** (Display system I/O resources).
5. Press **F6** to print the configuration.
6. Select either option **1** (132 characters wide) or **2** (80 characters wide). Press **Enter** to print to a spooled file.
7. Press **F12** to return to the Display Partition Information screen.
8. Select option **2** (Display partition processing configuration).
9. Press the **Print Screen** key to print the configuration to a spooled file.
10. Press **F10** (Display main storage information).
11. Press the **Print Screen** key to print the configuration to a spooled file.
12. Press **F10** (Display interactive information).
13. Press the **Print Screen** key to print the configuration to a spooled file.
14. Press **F10** (Display processor identifiers).
15. Press the **Print Screen** key to print the configuration to a spooled file.
16. Print the spooled files that were generated.

3.12 IPL Process

When the primary partition is IPLed, it will not start the IPL of the secondary partitions, unless specific criteria is entered. First, the serial number must be validated. Second, the system password checked to ensure it has not expired.

If the secondary partition was configured with a System IPL Action of *IPL*, the primary partition signals the secondary partition to start the IPL. The IPLs of all secondary partitions are started at the same time, and run concurrently. During this IPL process, all hardware checks are bypassed. Some of these were already completed by the IPL process of the primary partition. A normal IPL process of the secondary partition is significantly shorter than the primary partition.

3.13 Licensed Key Management

A licensed authorization code must be installed on your system. If the system prompts you to enter this key, then it will allow you to bypass the prompt and IPL the system for a maximum of 70 days. After the 70-day grace period, OS/400 will not IPL until a valid key is entered.

On a partitioned system, the licensed authorization code supplied by IBM has to be entered into each individual partition since each partition has its own copy of OS/400. The same license key can be entered for all logical partitions that are defined on the system.

There are two licensed keys. One is for the system password. The other is for OS/400 licensed users. Licensed keys for all other licensed programs and applications also have to be entered on *all* partitions. There is no automatic transfer of license information between the partitions.

3.14 Work Management

Take the following information into account when adjusting the work management setup for each partition.

3.14.1 Machine Pool

Every AS/400 system automatically reserves an amount of the machine pool for internal use. The system enforced minimum value varies depending on the main storage size of the machine. The total current reserved size can be seen when using the Work with System Status (WRKSYSSTS) command. When creating or changing a partition, the minimum and maximum values used for main storage can have a direct influence on the machine pool reserve size.

When the system is IPLed, it calculates the amount of main storage that is required to support the internal hardware page table for that maximum value. Then, the system adds this amount to the reserved figure. If the maximum value is set to an unrealistic value for that partition, the main storage is unnecessarily reserved in the machine pool and unavailable for use by other tasks and partitions.

For example, a system has 20 GB of main storage with three partitions. If each partition has the maximum amount of main storage set to 20 GB, the system reserves 512 MB in the machine pool for the hardware page table. It is obvious that it would not be possible for each partition to have 20 GB of main storage allocated at the same time. In this case, some of the 512 MB reserved in the machine pool is wasted.

At V4R4M0, the minimum and maximum values primarily affect when an IPL is required. If you change the amount of main storage allocated to a partition and it is within the minimum and maximum values, then only an IPL of that partition is required to make the change active. If the change to the amount of main storage allocated also requires a change to either of the minimum and

maximum values, then this requires an IPL of the primary partition to activate this change.

Note

The main storage totals seen on the WRKSYSSTS display may not equal the total for the logical partition. The difference is related to the size of the hardware page table. To reduce the difference, the maximum value for the partition should be lowered to be closer to the actual amount allocated.

3.14.2 QPFRADJ

The systems performance adjuster examines configuration related objects as one of the many inputs. It also makes performance adjustments accordingly.

We recommend that all obsolete configuration related objects be removed. This is especially necessary if the partitioned system was originally a single N-way machine, and the creation of partitions made a number of configuration objects obsolete in the primary partition.

Chapter 4. Change Management for a Logical Partition

A partitioned system may need to be changed. As long as your system supports the criteria for partitioning, you can change the setup of partitions inside these criteria. For more information about the criteria, see Chapter 2, “Planning Considerations” on page 17. For a detailed description on how to set up or do changes to your partitioned system, you can also refer to the articles found on the LPAR home page at: `http://www.as400.ibm.com/lpar`

- *Planning for and Setting Up Logical Partitions*
- *Managing Logical Partitions*
- *Backing up and Recovering Logical Partitions*

It is important to plan what you will do before you start. Taking the time to plan will save steps and avoid trouble during the partition setup. It is also important to understand that partitioning a system or changing the partition setup, using dedicated service tool (DST) or system service tool (SST), may seriously impact your system. Therefore, only specially dedicated persons should be allowed to do this. See "Controlling your Partitions with DST and SST" in the *Managing Logical Partitions* article for specific details on what DST and SST can do for logical partitions.

Be careful to read the LPAR planning articles before you make any changes to a partition configuration. Otherwise, this may cause you to perform an Initial Program Load on your system. It may also cause problems for normal work going on the system. Changing your partition setup is done from DST in your primary partition.

4.1 Managing Partitions

Changing logical partitions consists of several tasks. You may change:

- Logical partition processing resources
- Logical partition configuration

These steps are documented in detail in the *Planning for and Setting Up Logical Partitions* article.

4.2 Managing User Applications in a Logical Partition

Each partition is seen as a separate system in itself. There are no interfaces between the different partitions, other than what you have between separate applications in separate machines. That means normal communication setup.

The high-speed internal OptiConnect that can be defined in logically partitioned systems will, from an application point of view, be seen as if it is between separate boxes. See Chapter 7, “Inter-Partition Communication” on page 133.

Installing new user applications or maintaining an old one should not involve any special considerations working in a partition compared to a single system. Any application that you use now to maintain or distribute changes in applications to other systems, will work for maintaining or distributing changes in applications in single-system partitions.

4.2.1 Distributing Changes through Management Central

If you have changes that you want to distribute to other systems or partitions, you may use Management Central to do this. Management Central provides support for managing multiple AS/400 systems and is part of AS/400 Operations Navigator. AS/400 Operations Navigator allows you to manage your AS/400 server through a graphical interface on your Personal Computer (PC). AS/400 Operations Navigator is part of AS/400 Client Access Express, which is free with V4R4 of OS/400. When you install Client Access Express, Management Central will not be automatically installed. You have to select to install it. The interprocess communication between AS/400 systems (in this case, partitions) is done with TCP/IP.

On your source system, which is your selected partition, you can create packages with files and objects that can be distributed to target systems, or other partitions. You may also specify actions, or a command, to be executed on your target system when this package is successfully sent. This package may be scheduled to be sent later.

This means that you can:

- Distribute a batch input stream and run it on a remote system
- Distribute a set of programs and start your application
- Distribute a set of data files and run a program that uses that data

If you have partitions running the same applications, you may set up your applications to periodically send changes or updates to these partitions. The packages may consist of files and objects or they may be packed into savefiles. Management Central then takes care of distributing and handling these packages.

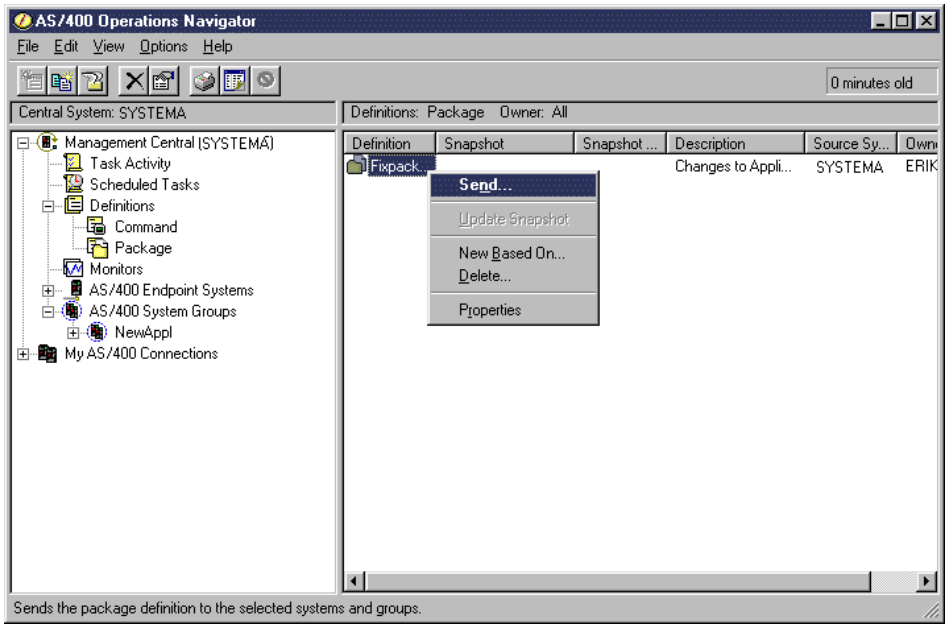


Figure 26. Sending a Package in Management Central

System Manager for AS/400 (5769-SM1) may also be used for distributing fixes between partitions. Independent software providers may use this product to distribute application fixes in the same way as normal OS/400 fixes. System Manager for AS/400 uses SNA for communication.

4.3 Program Temporary Fix (PTF) Management

The primary partition normally has the Electronic Customer Support line installed on the multifunction input/output processor (MFIOP). This is used for ordering PTFs for the primary partition. If communication IOPs are installed in the different partitions, you may use communication lines to define Electronic Customer Support lines for each partition.

If there are no communication lines, you have to copy the PTFs to media from the primary partition. Then, you must install it on the secondary partition from the saved media.

When installing PTFs, follow the steps that are described in "Installing temporary fixes on a partitioned system" in the *Managing Logical Partition* article.

4.3.1 Using Management Central in PTF Distribution

If you defined TCP/IP in the different partitions, you may use Management Central to deal with the fixes. Your primary system may be specified as your central system for Management Central. By using this as your endpoint system, you can copy PTFs from such media as tape or CD-ROM to your system.

In Configuration Services, you work with fixes on your system. Here you may install fixes for all or specific products. You may install them at once or schedule when you want to install them.

To distribute the PTFs to a partitioned system, keep the fixes in savefiles on your primary partition. You can use Management Central to copy the fixes from media, or order the fixes through the Electronic Customer Support line. Using Management Central, perform the following steps:

1. Right click on your system under Endpoint Systems.
2. Click on **Fixes—>Compare and Update** (Figure 27 on page 67).

   Now, you are in the wizard for setting up the transfer of fixes to other systems, which, here, is partitions.

3. Select the model system that is your primary system.
4. Select available systems or groups that are your partitions.
5. Select each of the products for which you are choosing to work with fixes.
6. Perform one of the following tasks:
   - Compare the results. Compare fixes on the primary partition against the secondary partition to see which fixes are missing on each secondary partition.
   - Send missing fixes from the primary partition to the secondary partition that you have chosen.
   - Send and install missing fixes on the secondary systems.
7. Run what you have chosen now or schedule it for later. If the fixes need an Initial Program Load, you may specify it now.

Putting on new cumulative packages with fixes is not handled through Management Central. Complete the instructions following the cumulative packages. You should install the cumulative packages using the Program Temporary Fix menu. This means you should use the CD-ROM installation media to install these package fixes. Since Management Central uses savefiles to distribute fixes, you may copy single fixes from the cumulative fix media. Then, install the fixes using Management Central.

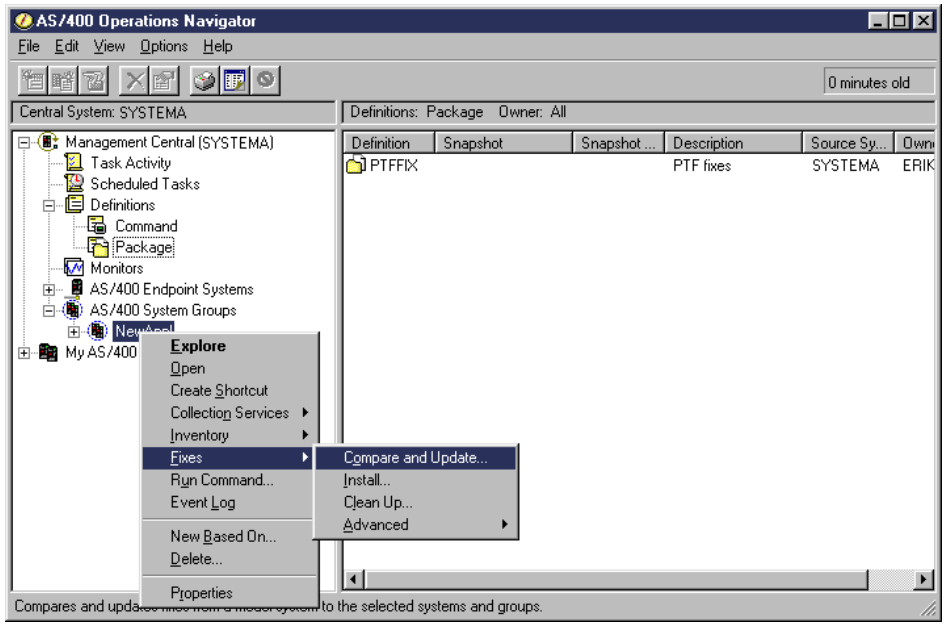


Figure 27. Sending Fixes Using Management Central

4.3.1.1 Why Manage Fixes with Management Central

One of the key benefits of Management Central is that it makes managing multiple systems, in this case partitions, as easy as managing a single system. Managing fixes is simplified with the tools that are included with Management Central. These tools include several wizards to guide you through these tasks:

- Installing fixes
- Permanently installing fixes
- Sending and installing fixes
- Uninstalling fixes
- Comparing and updating fixes

For example, to install multiple fixes, you select the fixes from a list and start the install wizard. You may find the Compare and Update wizard to be very beneficial. The wizard compares the fix level of a single system or multiple systems to a model system and makes them the same. Using the collected inventory, you can have the system compare the fix levels for a group of systems to a model system, send the save files of the selected fixes from the source system, and install the fixes.

4.4 Managing Security in Logical Partition

The security-related tasks you perform on a partitioned system are the same as on a system without logical partitions. However, when you create logical partitions, you work with more than one independent system. Therefore, you have to perform the same tasks on each logical partition instead of just once on a system without logical partitions. Here are some basic rules to remember when dealing with security on logical partitions:

- Add users to the system one logical partition at a time. You need to add your users to each logical partition that you want them to access.
- Limit the number of people who have authority to go to Dedicated Service Tools and System Service Tools on the primary partition. Refer to the table in "Controlling your partitions with DST and SST" in the *Managing Logical Partitions* article for more information on DST and SST.
- Secondary partitions cannot see or use main storage and disk units of another logical partition.
- Secondary partitions can only see their own hardware resources.
- The primary partition can see all system hardware resources in the Work with System Partitions displays of Dedicated Service Tools and System Service Tools.
- The operating system still only sees resources available to its own partition.

To manage users and user groups in your partitions, you can use Operations Navigator in AS/400 Client Access Express. From Users and Groups option, you can see all users and groups on an AS/400 system. From there you can manage AS/400 users and groups. For each AS/400 system to which you are connected, you can work with all users, groups, and users not in a group. You can perform the following actions with Users and Groups:

- Create and change users
- Add users to groups
- Copy users to other AS/400 systems (partitions)
- Delete users
- Create or change a group profile
- Drag and drop users and groups

  For example, you can drag a user to Users Not in a Group to remove the user from all groups. You can also drag a user to a group name to add the user to the group, add a user to another system (partition), or move the user to the group.

Chapter 5. Problem Management for a Logical Partition

If you encounter problems on a partitioned AS/400 system, you should first try to determine if they are general AS/400 problems, or problems specific to logical partitions. Refer to your AS/400 manuals on problem determination to evaluate whether your problem is a general AS/400 problem.

For recovery information in regard to logical partitions, refer to the *Backing Up and Recovering Logical Partition* article. This article is available from the LPAR home page at: `http://www.as400.ibm.com/lpar`

For other partition specific problems, you can try any of these options:

- Work with logical partition error messages and reports that appear on your console.
- Work with main storage dumps.
- Use remote service to get service on your system through a modem.
- Find out when you should power off and on a domain.
- Find out when you need to reset a disk unit input/output processor.

These options are also documented in "Troubleshooting logical partitions" in the *Managing Logical Partitions* article (also available from the LPAR home page).

5.1 Problem Determination for a Logical Partition

If the system experiences logical partition errors, it notifies you in one of two ways:

- Through an error message that appears at the bottom of the display
- Through the Logical Partitioning Error Report display that appears on the system console

From the primary system console, you may look at System Reference Codes and Initial Program Load codes for all partitions. From the secondary partition console, you may look at that partition's System Reference Codes and Initial Program Load codes for secondary only. You may choose to only see the main codes or use a command key to display all codes. See the "Displaying reference code history for secondary partitions" in the *Managing Logical Partitions* article.

5.2 Reporting to IBM

Each partition is like a single system. You may report problems from the problem log. This means that there has to be an Electronic Customer Support line defined in the partition. Through communication, you may set up one partition to work as a "hub" for the other partitions.

The Service Director is a tool that runs on the AS/400 system to supervise any problems encountered. It may send for Program Temporary Fixes or send messages to IBM informing of problems in the system. The main purpose is to encounter and handle problems before they get into serious problems for the system. The Service Director runs in each partition as if it is a single system and uses an Electronic Customer Support line for the partition in which it runs. Through communication lines between partitions, the Service Director may be set up to use the Electronic Customer Support line in one partition to communicate to IBM.

5.3 System Reference Codes (SRC)

The control panel on the system unit will show System Reference Codes (SRC) for the system primary partition only. If no partitions are defined, the entire system is viewed as a partition.

Each partition has its own Product Activity Log (PAL) and Service Action Log (SAL) accessed through dedicated service tools (DST) or system service tools (SST) that contain SRCs. The problem log also contains problem determination information that you can access by using the Work with Problems (`WRKPRB`) command.

Up to 200 of the most recent System Reference Codes for secondary partitions can be found by using the primary partition's console and accessing either DST or SST. In addition, you must perform the following steps:

1. Enter the `STRSST` command on the command line.
2. Select option **5** (Work with system partitions).
3. Select option **1** (Display partition information).
4. Select option **7** (Display secondary partition reference code history).

Or use the following steps to gain access through DST:

1. Force dedicated service tools on the primary partition.
2. Sign on to DST.
3. Select option **11** (Work with system partitions).
4. Select option **1** (Display partition information).

5. Select option **7** (Display secondary partition reference code history).

See Figure 28 for an example of the System Reference Code listing.

```
Secondary partition(s) to display . . . . .  *ALL  *ALL, 1-1
Number of reference codes to display  . . .  200   1-200

Partition
Identifier  Name      Reference Codes  Date      Time
  1         LPAR2     11               01/26/99  11:03:33
                      11 C900 2F00     01/26/99  11:03:33
                      11 C900 2C25     01/26/99  11:03:33
                      11 C900 2C20     01/26/99  11:03:32
                      11 C900 2C40     01/26/99  11:03:32
                      11 C900 2C10     01/26/99  11:03:00
                      11 C900 2B40     01/26/99  11:02:59
                      11 C900 2B30     01/26/99  11:02:58
                      11 C900 2B10     01/26/99  11:02:48
                      11 C900 2AC0     01/26/99  11:02:48
                      11 C900 2AB0     01/26/99  11:02:48
                      11 C900 2AA5     01/26/99  11:02:46
                      11 C900 2AA4     01/26/99  11:02:46

F3=Exit   F5=Refresh               F6=Print
F9=Include reference code detail   F12=Cancel
```

Figure 28. Display Secondary Partition Reference Code History

The reference code history log lists the SRCs. If the F9 (Include reference code detail) key is used and extended, the reference codes 12 through 19 are shown. See Figure 29.

```
Partition
Identifier  Name      Reference Codes  Date      Time
  1         LPAR2     11               01/26/99  11:03:33
                      11 C900 2F00     01/26/99  11:03:33
                      12 0DB0 0060
                      13 0000 0000
                      14 0000 0000
                      15 0000 0000
                      16 0000 0000
                      17 0000 0000
                      18 0000 0000
                      19 0000 0000
                      11 C900 2C25     01/26/99  11:03:33
                      12 0DB0 0060
                      13 0000 0000

F3=Exit   F5=Refresh               F6=Print
F9=Exclude reference code detail   F12=Cancel
```

Figure 29. Include Reference Code Detail

5.3.1 Licensed Internal Code Log

The Licensed Internal Code log is accessed under Dedicated Service Tools or System Service Tools. It has listings of major and minor codes that can be used to compare with the Product Activity Log entries. Many times, PTFs are written to fix problems that generate System Reference Codes or message IDs that occur with specific Licensed Internal Code log entries. The latest PTF for your release called the Preventive Service Planning (PSP), can be downloaded electronically to your system using the Electronic Customer Support line. Use the Send PTF Order command as shown here:

```
SNDPTFORD SF98VRM
```

Here, `VRM` is the version, release, and modification of OS/400 that is on your system. If a PTF is found within the Preventive Service Planning to fix the problem you encountered, follow the normal PTF procedures to procure, load, and apply the fix.

5.4 Power Interruptions in the Logical Partition

Power interruptions are detected by the primary partition. The primary partition then notifies the secondary partitions. This is an internal process that needs no setup from the user. Each partition receives the message about power interruption in the message queue specified in QUPSMSGQ. The partition has to take actions as if it is a single system.

Note: The system value QUPSDLYTIM is a common system value for all partitions.

5.5 Electronic Customer Support Lines

Only one Electronic Customer Support line is recommended per system. However, if a partition will be used independently, it may be desirable for each partition to have a separate service connection to order and receive PTFs.

There are three choices:

- A separate Electronic Customer Support communication line for each partition.
- One Electronic Customer Support communication line for the primary partition, which will act as a "hub" for the secondary partitions. Use

Management Central in Operations Navigator in AS/400 Client Access Express to distribute the fixes. See Chapter 4, “Change Management for a Logical Partition” on page 63, about PTFs.

- One Electronic Customer Support communication line set up to be switchable between secondary partitions.

In some places, it may be possible to receive PTFs over the Internet.

5.6 Logical Partition Configurations

Configuration data for all logical partitions is maintained on the load source disk of each logical partition. You cannot save the configuration data to tape. If a load source disk is initialized, the configuration data is copied from another load source disk during the Initial Program Load. See the *Backing up and Recovering Logical Partitions* article for recovery and working with configuration data for logical partitions.

For partition configuration information, see "Displaying and printing logical partition information" in the *Managing Logical Partitions* article.

5.6.1 Working with Configuration Data for Logical Partitions

The configuration data maintained on the primary partition is considered the master copy. Problems can occur when configuration information on the primary partition and a secondary partition conflict, or after you initialize disk units.

The Recover Configuration Data menu (see Figure 30 on page 74) in DST allows you to perform a different type of recovery on logical configuration data.

```
                        Recover Configuration Data
                                                              System: SYSTEMA
Select one of the following:

     1. Recover primary partition configuration data
     2. Update configuration data
     3. Clear non-configured disk unit configuration data
     4. Clear non-reporting logical partitioning resources
     5. Accept load source disk unit
     6. Copy configuration data to other side
     7. Clear configuration data









Selection


F3=Exit    F12=Cancel
```

Figure 30. Recover Configuration Data

5.6.1.1 Recovering Primary Partition Configuration Data

Recovering primary partition configuration data applies only to primary partitions and is part of a full system recovery. For more information, see the *Backup and Recovery* manual, SC41-5304. After re-installing the Licensed Internal Code on the primary partition, stop at the Initial Program Load. Or, install the system display and go to Dedicated Service Tools. In the Work with System Partitions menu, you will find Recover Configuration Data.

Select option **1**. Doing so locates a nonconfigured disk unit that contains the most recent configuration data for your system. The message `No units found with more current configuration data` may display instead of a list of resources. In this case, no unassigned disk units contain any appropriate configuration information.

You should be able to select a resource with a date and time that fall within a valid time frame for the configuration. When you confirm it, the system copies the new information data to the primary partition load source and automatically restarts the primary partition. The next time you perform an Initial Program Load on your secondary partitions, the system updates their logical partition configuration data from the primary partition.

5.6.1.2 Updating Configuration Data

You should update configuration data only when instructed to do so by a service representative. You can select the option to update the configuration data for every logical partition by copying it from the primary partition to all active logical partitions.

Important

Using this function when the configuration data on the primary partition is incorrect can ruin your existing configuration.

5.6.1.3 Clearing Nonconfigured Disk Unit Configuration Data

When you move disk units between logical partitions or systems, you may need to erase any old configuration data before the system can use the logical partition again.

Clear the configuration data when disk units have these characteristics:

- A System Reference Code of B600 5311 is reported in the product activity log against a nonconfigured disk unit.
- It is no longer a load source in its own logical partition.
- It is originally from a different logical partition where it was a load source, or it is from a different system which had logical partitions.

Selecting this option takes effect immediately.

5.6.1.4 Clearing Nonreporting Logical Partitioning Resources

This procedure should *not* be run if *any* hardware is marked as "failed". It should only be run when all system hardware is completely operational.

After you add, remove, or move hardware within a system with logical partitions, you may have resources that are no longer available or that are listed twice. Old entries may still exist in the configuration data. You can clean up these listings so that all nonconfigured resources are erased from the logical partition. You do this from Dedicated Service Tools display.

5.6.1.5 Accepting a Load Source Disk Unit

When configuration data on the load source of a logical partition is different from what the system accepts, an error is logged in the Product Activity Log with the System Reference Code B600 5311.

If you recently moved or added disk units either within the system or from another system, they may still contain configuration data. If you do not want to use this disk unit as the load source, you need to clear its configuration data before proceeding. Otherwise, you may want to use this disk unit as your new load source on the logical partition. In this case, you should select this option from a DST display.

5.6.1.6 Copying Configuration Data to the Other Side

Your system may experience a disk read error of the logical partition configuration data if the following reference codes appear when you restart from one source but not another:

- 11 B193 4511
- 12 xxxx xx5D (where x equals any value 0 through 9 or A through F)
- 13 690A 2060

You can find these reference codes on the operations panel. Copy the data from the functioning source to the faulty source with this procedure.

Note: Only attempt to do this when you are certain that the logical partition restarts normally when using the other Initial Program Load source.

Ensure that the configuration information contained on this Initial Program Load source is correct. Verify this by viewing your logical partition's status. A correct configuration would show your most recent configuration of logical partitions. If it is correct, start copying the configuration data. If the configuration is not correct, do not copy the data.

This option writes the logical partitioning configuration data stored on this side of the load source disk unit to the other side. For example, it can write from source A Licensed Internal Code to source B Licensed Internal Code, or from source B Licensed Internal Code to source A Licensed Internal Code.

5.6.1.7 Clearing Configuration Data

If you need to clear all configuration data to return the system to a nonpartitioned system again, use the clear configuration data option. In doing so, the system will no longer be partitioned.

This option returns all hardware resources to the primary partition. However, all user data within secondary partitions will be lost. Be careful and perform adequate backups before selecting this option.

Restart the system in manual mode for the change to take effect. On the Initial Program Load, go to Dedicated Service Tools and select the option to clear configuration data from nonconfigured disk units on the Recovery

Configuration Data display. Do this once for every former secondary partition load source disk unit.

5.7 Recovery for Logical Partitions

Recovery for logical partitions falls within two main categories:

- Recovery of logical partition configuration data (this must be done one logical partition at a time)
- Recovery of user data to a logical partition

This section explains these categories.

5.7.1 Recovery Considerations with Logical Partitions

As you go through the recovery steps outlined in the *Backup and Recovery* manual*,* SC41-5304, keep these points in mind:

- Recovery occurs on a partition-by-partition basis. There is not a system-wide recovery option for all logical partitions.
- Since each logical partition should have its own backup media, ensure that you use the appropriate media for the logical partition you recover.
- Make sure that you recover the logical partition configuration data after initializing the disk units of the primary partition.
- On a full system recovery, you must perform the following steps:

  1. Restore the Licensed Internal Code to the primary partition.
  2. Recover the configuration data for the logical partitions.
  3. Recover the disk configuration for the primary partition.
  4. Install the operating system on the primary partition.
  5. Recover any of the secondary partitions by restoring the Licensed Internal Code, recovering the disk configuration, and installing the operating system.

  These steps are covered in more detail in the "Recovering Information on Your System" section in the *Backup and Recovery* manual*,* SC41-5304.
- You work with configuration data for logical partitions through a menu in Dedicated Service Tools. The configuration data for the logical partitions is never restored from your backup media.
- For secondary partitions that are sharing devices, make sure that you switched the device to the required logical partition before trying to perform a recovery.

- You can view System Reference Code activity for secondary partitions through the primary partition in Dedicated Service Tools or System Service Tools. You can also view System Reference Codes activity for a secondary partition from its own Dedicated Service Tools display or its own System Service Tools display. You monitor the System Reference Code on the Display Partition Status screen.
- Look for all configuration data errors that are related to logical partitions in your Licensed Internal Code log and Product Activity Log. Each logical partition also has its own logs that you can view. However, some errors on a secondary partition can only be found in the logs on the primary partition. Be sure to check both places when possible.

5.7.2 Information on Recovering after System Initialization

Anytime you initialize the disk units on the primary partition, you need to recover your configuration data. Initialization can occur just prior to reinstalling the operating system. You need to recover the configuration because initializing the disk units erases everything on them, including user data and configuration data.

When you recover your configuration, the system looks for the configuration data that is stored on the secondary partitions. Then, it copies the configuration data to the primary partition. When you erase the disk units in a secondary partition, the configuration data is automatically copied from the primary partition to that secondary partition once you restart it.

First, you re-install the Licensed Internal Code on the primary partition. Immediately after the system restarts and before you configure disk units or recover Auxiliary Storage Pool disks, follow the steps to recover your configuration as shown in Figure 40 on page 101.

5.7.3 Restoring User Data to a Logical Partition

Restoring user data to a logical partition follows most of the same steps as restoring data to a nonpartitioned system. However, there are a few considerations to keep in mind before you begin recovery procedures. See the "Recovering Information on Your System" section in the *Backup and Recovery* manual, SC41-5304, for the specific instructions on data recovery for your system.

5.7.4 Recovering User Data from a Deleted Logical Partition

In general, you cannot recover user data or system data from a deleted logical partition. However, if a logical partition is accidentally deleted, it may

be possible to manually recreate it without any data loss. Disk units from the deleted partition must not meet any of the following conditions:

- Moved to another logical partition
- Removed from the system
- Configured into another logical partition or system

To recover the user data and system data, perform the following steps:

1. Create a new logical partition.
2. Add *all* resources back into the deleted logical partition so that it has the same configuration as it did before it was deleted.
3. When you finish creating the logical partition, restart the system in normal B mode.
4. Recover the user data and system data from the disk units.

5.8 Service Provider Issues

Servicing a partitioned system is like servicing multiple separate systems. A failure in a secondary partition would allow the other partitions to continue operations. However, there are certain failures that are not partition dependent that cause the *entire* system to halt:

- Processor failure
- Main storage failure
- Service processor (MFIOP)
- Power
- Any time the primary partition is down, and then all partitions are down!

There are questions that come up when servicing, such as where is my console, load source, or alternate Initial Program Load device for the secondary partitions. Since they can physically be located in different frames, it is important to do the following labeling or have printouts available:

- Rack configuration and partition configuration
- Consoles labeled for each partition
- Label racks if bus-level partitioned

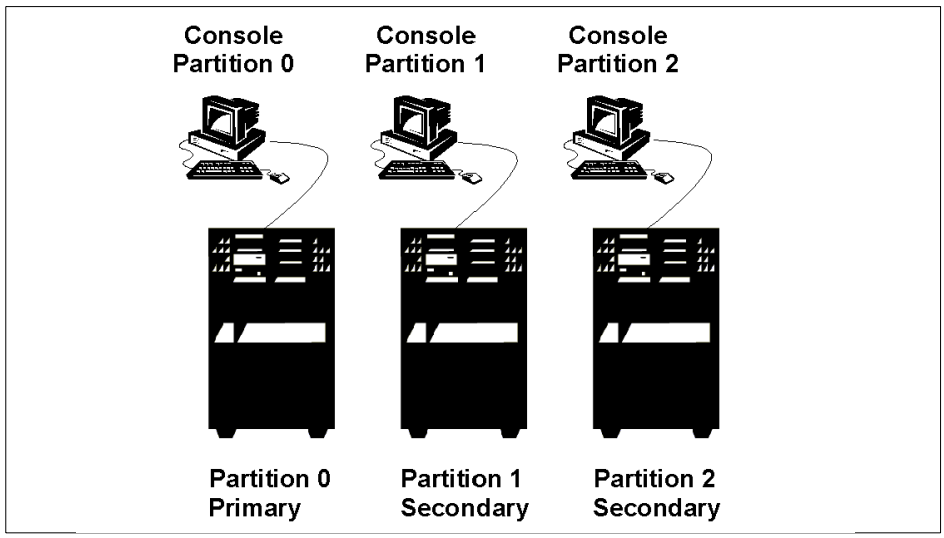


Figure 31. System Properly Labeled

Concurrent maintenance on a device must be done within the partition to which the input/output processor is dedicated.

Chapter 6. Capacity Planning and Performance

In Chapter 1, “Logical Partitioning Concepts” on page 1, we learned that logical partitioning can be extremely useful in fulfilling a wide range of customer needs. The success or failure of an implementation of logical partitioning in any situation depends largely upon an appropriate analysis of the situation’s requirements.

6.1 Determining Customer Requirements

The basic questions to be asked before contemplating any kind of solution are:

- What is the customer trying to achieve?
- How is the customer’s business structured?
- What is the best solution possible for their particular situation?

Consider the following example:

You are approached by a manufacturing company that wants to consolidate its worldwide operations. Currently, it has five branches in five different countries. Each branch has its own IT division, that runs the same applications, and transfers the results of the day’s activities daily to the machine at headquarters. Here, all results are consolidated, manufacturing schedules are established, and consolidated reports are compiled. The CEO has decided that:

- One daily transfer of the day’s activities does not provide him with enough control of the company’s manufacturing schedules and current financial situation. He needs instant access to up-to-date information.
- Running five IT divisions is becoming too costly.
- The company needs to consolidate its IT divisions into one at headquarters.

There are a number of solutions to this unlikely situation. For purposes of discussion, let us assume that you decide to offer a logical partitioning solution. You now need to go through some or all of the following steps:

1. Analyze the workload on each physical machine (maybe using BEST/1).
2. Determine the amount of interactive and batch resources needed.
3. Determine languages to be used.
4. Determine the appropriate communications requirements.
5. Understand service level agreements and problem or change management disciplines.

Up to this point, you followed the standard steps towards any solution. From now on, your focus is narrowed down to logical partitioning issues.

You now need to perform some capacity modelling to determine the batch and interactive resources required per partition. This will help determine what type of physical machine is needed. BEST/1 is a useful tool in this respect. In 6.6.2, “BEST/1” on page 99, we go into the finer details of a workable methodology to determine the resources required using BEST/1. For now, we look at the bare mechanics of the process.

Decide the machine model, number of partitions, batch or interactive resources required per partition, number of processors, and amount of main storage per partition. Now, you are ready to enter the next phase of this process: Planning for the full hardware configuration. Refer to Chapter 2, “Planning Considerations” on page 17, for the procedures to follow.

The following sections describe the steps for the actual allocation of resources to create a multi-partitioned machine for our hypothetical company.

6.2 Commercial Processing Workload (CPW) Concepts

Commercial Processing Workload (CPW) represents a measure of relative system performance. Different AS/400 models have different CPW ratings. Traditional AS/400 systems have a single CPW value, which represents the maximum workload for this model. This CPW value is valid for an interactive workload, a client/server workload, or a combination of the two workloads.

6.2.1 AS/400 Workload — Technical View

From a technical point of view, the AS/400 workload is divided into two major sections:

- **Interactive workload**

  The interactive workload or job allows users to see a 5250 display. It can be a non-programmable terminal (dumb terminal), PC with emulation software, or an Internet or intranet function with Workstation Gateway. It usually has a priority level of 20.

- **Non-interactive workload**

  In a non-interactive workload, users do not see a 5250 display. An example of this workload is a PC program that accesses the AS/400 database using a client/server technique and submitting batch jobs. AS/400 system jobs are also non-interactive types of work. Most of the

non-interactive jobs have a run priority greater than level 20. Some of them, such as system jobs, have a priority less than level 20

Note: 0 is the highest priority level, and 99 is the lowest priority level.

6.2.2 Different CPW Values

On a server model, there are two CPW values. The larger CPW value shows the maximum workload for that model if only client/server applications (no interactive tasks) were running. The smaller number shows the maximum workload for that model if only interactive applications were running. These two values are not *additive*. This means that interactive CPW and client/server CPW cannot be added together since interactive CPW is part of the client/server CPW value. Interactive tasks reduce the client/server processing capability of the system.

Custom server models act similar to server models. But, custom servers follow a different algorithm. Therefore, the "knee" of the curve for workload interaction is at a different point.

The AS/400 server 7xx models (announced February 1999) and the new Model 170 servers (announced September 1998) also have two CPW values. The higher value shows the relative processor capacity. Therefore, there is no need to differentiate between constrained and unconstrained CPW values. The smaller CPW value represents the maximum interactive workload for the specific model. If you add an interactive workload to your Model 170, the remaining client/server capacity available is calculated as:

```
CPW (C/S batch) = CPW (processor) - CPW (interactive)
```

In performing capacity planning for logical partitions, it is important to know what the specific algorithm for that AS/400 model is. On the new AS/400 7xx models, you can create a logical AS/400 partition that:

- Suits your own specific needs
- Works like a traditional AS/400 system model
- Works like a custom server model
- Works like an AS/400 server model

6.2.3 AS/400 System Models

For traditional AS/400 system models, the full CPU processor capabilities are available to either standalone interactive, stand-alone non-interactive, or mixed-mode work environments. The impact between is are controlled by the job Run Priority, Time Slice End, and Storage Pool parameters.

Figure 32 shows how system performance is on an AS/400 Model 600 if you mix interactive and client/server (batch) workloads.

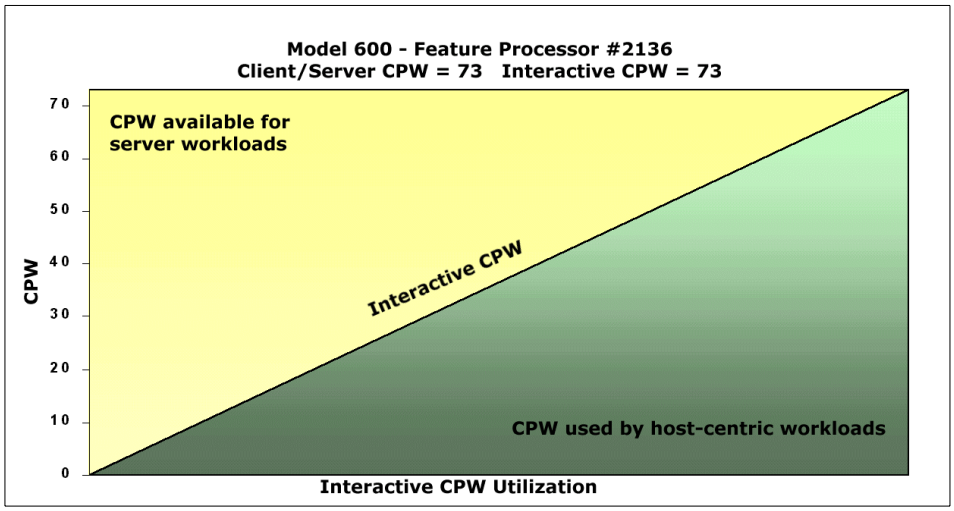


Figure 32. Interactive CPW Utilization — Model 600

Table 3 offers another way to view performance based on specific interactive percentages.

Table 3. AS/400 Model 600 — Feature Processor #2136

Interactive Percentage	0%	33%	66%	100%
CPW interactive	0.0	24.1	48.2	73
CPW client/server	73	48.7	24.3	0

6.2.4 New AS/400e server 7XX

Only for the new February 1999 7XX processors and the September 1998 170 processors, the system will manage the interactive work differently than on other servers. In an interactive-only environment, the interactive jobs (5250 green screen jobs) will get the rated CPW (or slightly higher CPW rating). However, the CPU may not be 100% utilized. You will still find that the non-interactive-only environment gets the full non-interactive CPW, now called Processor CPW rating. In an environment of both interactive and non-interactive work, the interactive work will be permitted to reach its CPW rating, but not any higher.

The new 7XX and 170 models have interactive and processor CPW ratings. The 170 models have a fixed interactive rating. For the new 7XX models, you

can choose the interactive CPW you need for your workload. In a mixed interactive and non-interactive workload environment, you can size the system based on subtracting interactive CPW from the full non-interactive - processor CPW rating.

For example, a Model 730-2068 has a processor CPW rating of 2,890 and an interactive CPW rating (with interactive feature 1511) of 2,000. When the interactive jobs use 2,000 CPW (actually the system manages this as a percentage of CPU utilization), no more interactive job throughput can be achieved. However, at that level, there is CPU left for non-interactive jobs to achieve 890 CPW amount of work.

Figure 33 shows how system performance is on an AS/400 Model 730 if you mix interactive and client/server (batch) workloads.

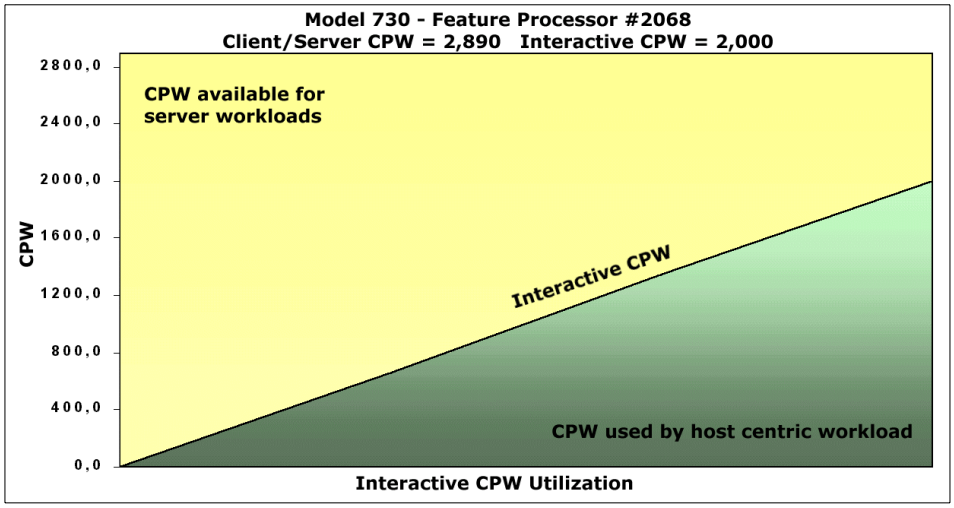


Figure 33. Interactive CPW Utilization — Model 730

Table 4 offers another way to view performance based on specific interactive percentages.

Table 4. AS/400 Model 730 — Feature Processor #2068

Interactive Percentage	0%	33%	66%	100%
CPW interactive	0.0	660	1320	2000
CPW client/server	2890	2230	1570	890

6.2.5 AS/400 Server Models

For server models, the full CPU processor capabilities are available only to non-interactive work environments. The non-interactive CPW rating is achieved in a non-interactive only job environment at 100% CPU utilization. For the interactive CPW rating, 5250 work consumes 100% CPU utilization, but cannot make full use of the maximum CPU power.

At approximately one third of the interactive CPW rating, the interactive work begins to degrade any concurrent non-interactive work. The impact accelerates dramatically as interactive work grows above this one-third CPW value.

Figure 34 shows how system performance is on an AS/400 Model S30 with processor feature 2259 if you mix interactive and client/server (batch) workloads.

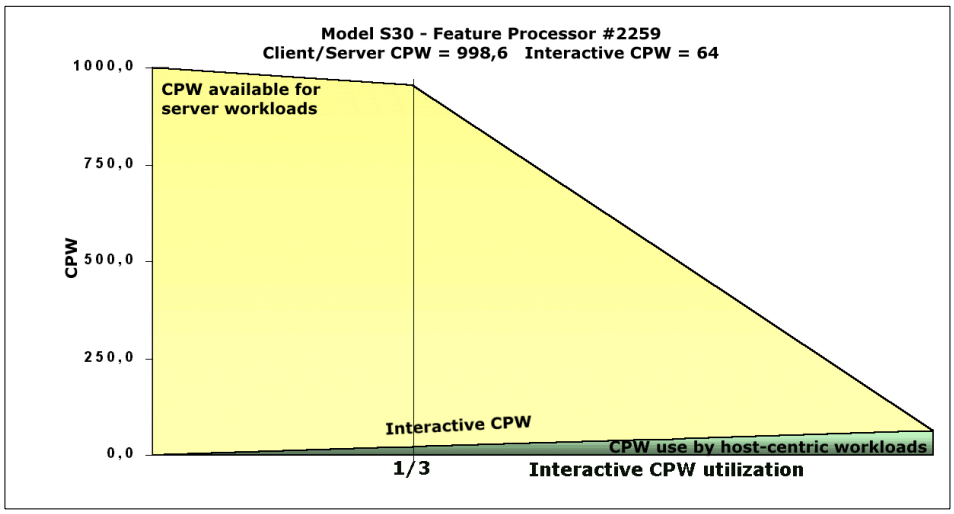


Figure 34. Interactive CPW Utilization — Model S30 Processor #2259

Table 5 offers another way to view performance based on specific interactive percentages.

Table 5. AS/400 Model S30 — Feature Processor #2259

Interactive Percentage	0%	33%	66%	100%
CPW interactive	0.0	21.1	42.2	64.0
CPW client/server	998.6	931.9	466.0	0.0

6.2.6 AS/400 Custom Server Models

For custom mixed-mode server systems, the full CPU processor capabilities are available only to non-interactive work environments. The non-interactive CPW rating is achieved in a non-interactive only job environment at 100% CPU utilization. For the interactive CPW rating, 5250 work consumes 100% CPU utilization, but cannot make full use of the raw CPU power.

The relatively higher interactive CPW rating on Custom Mixed Mode servers is achieved by the system permitting interactive work to grow until it is at six sevenths of the interactive CPW rating. Above this six-sevenths CPW rating, the interactive work impact on non-interactive work dramatically increases.

Figure 35 shows how system performance will be on an AS/400 Model S20 if you mix interactive and client/server (batch) workloads.

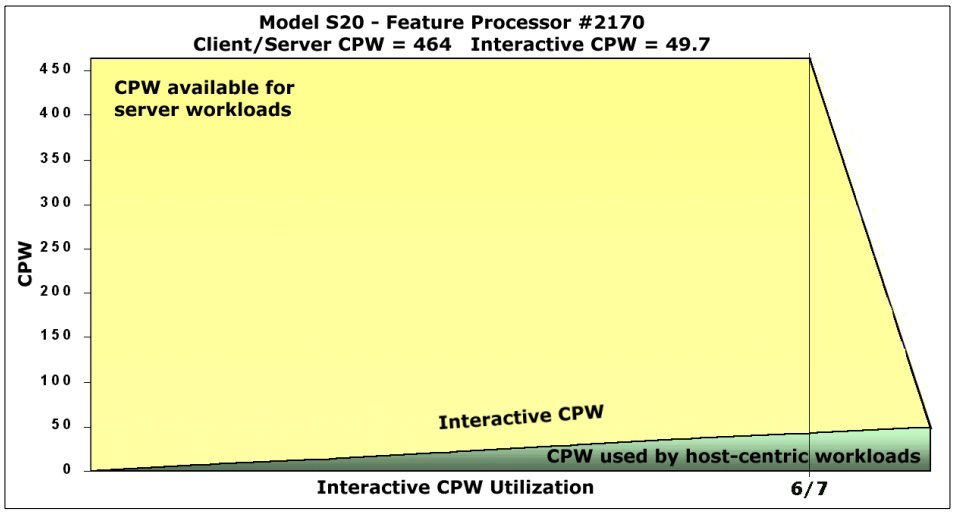


Figure 35. Interactive CPW Utilization — Model S20

Table 6 offers another way to view performance based on specific interactive percentages.

Table 6. AS/400 Model S20 — Feature Processor #2170

Interactive Percentage	0%	33%	66%	100%
CPW interactive	0.0	16.4	32.8	49.7
CPW client/server	464.3	447.4	430.9	0.0

6.3 Determining Batch CPW

Non-interactive workload includes the typical batch job as well as the AS/400 client/server workload. It is triggered by the Submit Job command, spooling, message queue, and data queue entries.

Typically batch functions include:

- Networking
  - APPC
  - TCP/IP
- Database
  - SQL
  - DDM
  - Remote SQL
  - ODBC
  - File transfer
- File serving
  - Shared folders
- Client/server applications
  - Progress/400
  - SYNON CSG
  - SQL Windows
  - ENVY/400
  - And so on
- Traditional batch
  - Development
  - Data warehouse
  - Internet/intranet
  - Domino/Notes

A non-interactive transaction includes:

- Job run time
- Time to complete printing a spooled file
- Time to complete a set of related jobs
- Time to process "n" message queue or data queue entries
- The number of records processed per hour
- Time to send or receive a file or files between a server and its attached client (or clients)

The performance objectives should show the amount of work performed and the amount of resources used in a specific amount of elapsed time. The amount of resources to perform this work is important as well since it defines the amount of CPU needed, the number of DASD arms, or the amount of DASD space needed.

Knowing the number of records processed per unit of time, the number of disk I/O operations per second, or the number of bytes (characters) transferred by these non-interactive jobs is, in many cases, the only way to set realistic expectations or evaluate degradation or improvement in batch run times.

6.3.1 BATCH400 — Enabling the AS/400 Batch Window

BATCH400 is a new AS/400 tool to help analyze and predict batch job performance on the AS/400 system. BATCH400 addresses the often asked question: *What can I do to my system to meet my overnight batch runtime requirements (also known as the Batch Window)?*

The BATCH400 tool creates a "model" from AS/400 performance data. This model resides in a file in the target library. The tool can be asked to analyze the model and provide results for various "what-if" conditions. Individual batch job runtime and overall batch window runtimes are reported by this tool.

BATCH400 allows you to "view" the batch window in a form that shows job start and stop times and dependencies. This should allow you to locate times in the batch window when more efficient job scheduling can improve total system throughput. BATCH400 can also help predict improvements in throughput that results from such hardware upgrades as CPU and DISK.

Note: Batch400 is an IBM internal-use-only tool.

6.3.2 Batch Capacity Planning for LPAR

Before you start capacity planning, be sure you have the right hardware and software requirements fulfilled for your system. Only OS/400 Version 4 Release 4 (V4R4) supports logical partitioning. V4R4 is the oldest release supported in any partition.

Logical partitioning is only supported on the AS/400 systems shown in Table 7.

Table 7. Systems Supporting LPAR — Batch CPW

Model	2-Way	CPW Batch	4-Way	CPW Batch	8-Way	CPW Batch	12-Way	CPW Batch
620	#2182	464.3						
640	#2238	583.3	#2239	998.6				
650					#2240	1,794	#2243	2,340
					#2188	3,660	#2189	4,550
720	#2063	810	#2064	1,600				
730	#2066	1,050	#2067	2,000	#2068	2,890		
740					#2069	3,660	#2070	4,550
S20	#2165	464.3	#2166	759.0				
S20 ISV	#2170	464.3	#2177	759.0				
			#2178	759.0				
S30	#2258	583.3	#2259	998.6	#2260	1,794		
S30 ISV			#2320	998.6	#2321	1,794		
					#2322	1,794		
S40					#2256	1,794	#2261	2,340
					#2207	3,660	#2208	4,550
S40 ISV					#2340	3,660	#2341	4,550

To determine the right batch CPW value for your partitioned system, you have to know the CPW value for that system. In Table 7, you can find the CPW value available for the eligible systems. On an N-way system, each processor has a specific CPW value. Divide the total CPW value by the number of processors to calculate the individual CPW per processor. If the value for a single processor meets your requirements for batch workload, you can setup a logical partition with one processor only. If you need more CPW for your batch workload, you have to add one or more of the available processors to that specific partition.

6.3.2.1 Example: Assigning CPW for a Batch

Before you start assigning CPW value for logical partitions, we strongly recommend that you do capacity planning. In this example (see Table 8), capacity planning showed a CPW value of 550 required for your batch environment.

Table 8. Assigning CPW for a Batch Environment

AS/400 Model	No. of Processors Available	CPW Available	Required Batch CPW
9406-720 #2064	4	1,600	550

First, you need to calculate the CPW value for each processor:

```
CPW (single processor) = CPW (batch) / number of processors
CPW (single processor) = 1,600 / 4 = 400 CPW
```

The maximum CPW value you can get in a partition with one processor only is 400 CPW. In this example (see Figure 36), you need 550 CPW for your batch workload. You need at least two processors for your batch workload. Only whole processors can be assigned to a logical partition. During setup for a logical partition, you are also asked for the amount of main storage for that partition. To determine the required amount of main storage, you have to perform capacity planning using BEST/1.

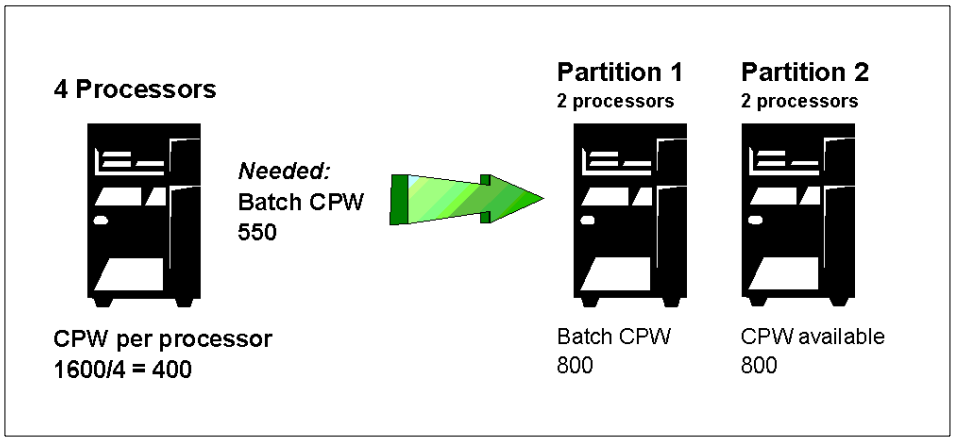


Figure 36. Assigning CPW for Batch

6.4 Determining Partition Processing Resources

In the process of setting up your partitions, you need to ensure that enough resources are available to provide the required machine performance. The following discussion attempts to show the pure mechanics of working out your logical partitioning configuration. This may not lead to a feasible configuration. Therefore, at a later time, we need to consider the pitfalls to avoid and refinements to apply in order to create a logically partitioned machine that can actually be physically implemented.

In the planning stages for your machine, you decide the overall scope of most of these parameters. Now, we drill down to the process of logically partitioning the machine by allocating the appropriate resources. Figure 37 shows the actual screen from which you will be working.

```
                        Create New Partition
                                                          System:   SYSTEMA
Complete blanks, press Enter.

  Partition identifier . . . . . . . . . . . . . .  0
  Partition name . . . . . . . . . . . . . . . . .

  Number of available system processors  . . . . :  8
  Number of partition processors . . . . . . . . .  1
  Minimum / maximum number of processors . . . . .  1    / 2

  Size of available system main storage (MB) . . :  24576
  Size of partition main storage (MB)  . . . . . .   2048
  Minimum / maximum size of main storage (MB)  . .   2048    /

  Interactive performance available  . . . . . . :  100 %
  Partition interactive performance  . . . . . . .   12 %
  Minimum / maximum interactive performance  . . .   12 /12   %

  Connect to system inter-partition OptiConnect  .     1=Yes, 2=No

F3=Exit                                          F9=Exclude limits
F11=Display partition processing configuration   F12=Cancel
```

Figure 37. Creating Logical Partitions

The three resources we need to consider now are processors, main storage, and interactive performance. We examine these items one at a time in the following sections.

6.4.1 Determining the Number of Processors

In your planning, before ordering your machine, you decide on:

- The number of partitions required (in our case, 5)
- The CPW requirements for each partition.

  These are obtained by modelling through BEST/1. For example, one environment produced a Model 720 with a processor CPW rating of 810. Refer to 6.2, “Commercial Processing Workload (CPW) Concepts” on page 82, for more details. You know that the partition to contain that environment needs this CPW rating. As an example, Table 9 provides details of the new 7XX models.

Table 9. CPW Ratings for 7XX Models

Model	Initial Proc. Code	# of Proc. #	Proc. CPW	Inter. Base	Additional Interactive CPW				
					1501	**1502**	**1503**	**1504**	**1505**
720	**2061**	1	240	35	70	120			
720	**2062**	1	420	35	70	120	240		
720	**2063**	2	810	35		120	240	560	
720	**2064**	4	1,600	35		120	240	560	1,050
					1507	**1508**	**1509**	**1510**	**1511**
730	**2065**	1	560	70	120	240	560		
730	**2066**	2	1050	70	120	240	560	1,050	
730	**2067**	4	2000	70		240	560	1,050	2,000
730	**2068**	8	2890	70		240	560	1,050	2,000
					1510	**1511**	**1512**	**1513**	
740	**2069**	8	3,660	120	1,050	2,000	3,660	4,550	
740	**2070**	12	4,550	120	1,050	2,000	3,660	4,550	

The complete capacity planning results for our hypothetical situation are summarized in Table 10.

Table 10. Capacity Planning Results

Environment 1	Environment 2	Environment 3	Environment 4	Environment 5
Model: 720 Feature: 2061 Prc.CPW: 240 Int. CPW :120	Model: 720 Feature: 2063 Prc.CPW: 810 Int. CPW :240	Model: 720 Feature: 2063 Prc.CPW: 810 Int. CPW :240	Model: 720 Feature: 2061 Prc.CPW: 240 Int. CPW :120	Model: 720 Feature: 2062 Prc.CPW: 420 Int. CPW :240

A possible mapping of the above into a logical partitioning context would appear as shown in Table 11.

Table 11. First Cut into Logical Partitions

Partition 0	Partition 1	Partition 2	Partition 3	Partition 4
Prc.CPW: 240 Int. CPW :120	Prc.CPW: 810 Int. CPW :240	Prc.CPW: 810 Int. CPW :240	Prc.CPW: 240 Int. CPW :120	Prc.CPW: 420 Int. CPW :240

This would have yielded a physical machine with a minimum of 2,520 processor CPW and 960 Interactive CPW. The nearest available machine matching these parameters is an 8-Way Model 730-2068 rated at 2,890 processor CPW and 1,050 Interactive CPW. With such a machine, the CPW per processor is 361 (2,890 divided by 8).

You can now work out the number of processors per partition as shown in Table 12.

Table 12. Working Out the Number of Processors

Partition 0	Partition 1	Partition 2	Partition 3	Partition 4
Prc.CPW: 240 Int. CPW :120	Prc.CPW: 810 Int. CPW :240	Prc.CPW: 810 Int. CPW :240	Prc.CPW: 240 Int. CPW :120	Prc.CPW: 420 Int. CPW :240
1 Processor 361 CPW	3 Processors 1,083 CPW	2 Processors 722 CPW	1 Processor 361 CPW	1 Processor 361 CPW

Notice that the processing power of a partition is tightly coupled with the number of processors it contains. For example, partition 1 needs three processors to fulfill its CPW quota. Only one processor is required for the primary partition (partition 0).

Looking at partition 2, we notice that our capacity plan specified 810 processor CPW, but we managed only 722 CPW. In a real situation, this would alert us to the fact that we need a physical processor with more CPW and need to go back to the drawing board. However, we will ignore this for this exercise.

The logical partitioning function allows you to specify a minimum and maximum range for the number of processors for each partition. This capability can be very useful in managing changes to a partition workload. For example, if, for some reason, there is a reduction in the workload in partition 1 and you find that it now needs only 600 CPW, you have the flexibility of reducing the number of processors, as long as you are within your defined range. The removed processors become available for an addition to any other partition needing more power, with minimum disruption.

Notes:

- Changes to processor resources beyond the defined range require a full system restart, a major disruption.
- The absolute minimum is one processor per partition. These minimums can be specified only if the maximum limit for the partition is the same.

6.4.2 Determining Main Storage

In planning user requirements, you would use a tool like BEST/1 to determine the amount of main storage needed. Table 13 shows how this would work.

Table 13. Working Out Memory Allocation

Partition 0	Partition 1	Partition 2	Partition 3	Partition 4
Prc.CPW: 240 Int. CPW :120	Prc.CPW: 810 Int. CPW :240	Prc.CPW: 810 Int. CPW :240	Prc.CPW: 240 Int. CPW :120	Prc.CPW: 420 Int. CPW :240
1 Processor 361 CPW	3 Processors 1,083 CPW	2 Processors 722 CPW	1 Processor 361 CPW	1 Processor 361 CPW
2,048 MB	**7,068 MB**	**4,096 MB**	**1,024 MB**	**3,192 MB**

You decide to add 35% more memory because you now have a good understanding of the customer's workload. This yields a Model 730 with 24,576 MB of main storage.

Using the screen pictured in Figure 37 on page 92, you can now assign main storage to each of your partitions. Again, specify the minimum and maximum range to allow some flexibility for on-the-fly main storage changes with minimum disruption.

Notes:

- Changes to main storage resources beyond the defined range require a full system restart, a major disruption.
- The absolute minimum main storage is 256 MB for the primary partition and 64 MB for the secondary partition.

6.4.3 Determining Interactive Performance

One of the most important conclusions you would reach in the initial analysis of your customer's environment is the type of workload in each environment, namely batch (client/server) or interactive (green screen). Your assessment would directly affect your choice of processor feature.

Refer to Table 10 on page 94 for the complete results of your capacity planning exercise.

You are now ready to assign interactive performance to each of your partitions.

Table 14. Working Out Interactive Performance

Partition 0	Partition 1	Partition 2	Partition 3	Partition 4
Prc.CPW: 240 Int. CPW :120	Prc.CPW: 810 Int. CPW :240	Prc.CPW: 810 Int. CPW :240	Prc.CPW: 240 Int. CPW :120	Prc.CPW: 420 Int. CPW :240
1 Processor 361 CPW	3 Processors 1,083 CPW	2 Processors 722 CPW	1 Processor 361 CPW	1 Processor 361 CPW
2,048 MB	7,068 MB	4,096 MB	1,024 MB	3,192 MB
Actual: 12% Min: 1% Max: 12%	Actual: 37% Min: 3% Max: 37%	Actual: 24% Min: 2% Max:24%	Actual: 12% Min: 1% Max: 12%	Actual: 12% Min: 1% Max: 12%

From a performance perspective, it is important to note that with the new 7XX models, the published interactive CPW is fully usable to 100% without incurring any performance penalty. This is different from prior models (for example, Model S30). In the earlier models, a performance penalty is incurred when interactive CPU usage reaches either one-third or six-sevenths of the published interactive CPW, depending on the model (Sxx or the newer mixed-mode servers).

When dealing with non-7XX server models, interactive performance calculations must take into account the fact that there is a drastic performance degradation if the CPU percentage of interactive usage exceeds the limits previously mentioned.

Note:

- Changes to interactive resources beyond the defined range require a full system restart, a major disruption.
- The absolute minimum interactive performance is 1% for each processor. This means, for example, that a three-processor partition will have a minimum interactive performance of 3% on a model 7XX.
- Notice that the maximum interactive performance that you can obtain depends on the number of processors in a partition. For example, in the first partition in Table 14 on page 96, the maximum interactive performance is 12%. This is the percentage of the interactive CPW for one processor out of eight. In the same way, in the second partition, which has three processors, the maximum interactive performance possible is 37% (three-eighths of the total interactive performance).

Refer to Appendix E, “Interactive and Batch CPW Determination” on page 189, for a more detailed scenario of the process.

6.5 OptiConnect/400 Performance Considerations

OptiConnect/400 allows applications written to access databases locally, to access them remotely with simple changes to file descriptions and no changes to applications. It provides an optimized Distributed Data Management (DDM) solution that is capable of efficient, low-latency, high bandwidth communication between AS/400 logical partitions.

Before V4R4 and logical partitioning, OptiConnect/400 utilized a fiber-optic connected, shared I/O bus between AS/400 systems. With V4R4, OptiConnect/400 can use either the special hardware I/O processors or the virtual OptiConnect on a single N-way system with logical partitions.

The following conclusions are derived from laboratory benchmarks and from experience gained by working with IBM OptiConnect/400 customers:

- Each time that a *put* or *get* is issued, OptiConnect/400 uses the shared I/O bus to exchange data with the remote system. DDM uses a communication link for accessing data on a remote system. In fact, this is different if you use virtual OptiConnect on logical partitions. Virtual OptiConnect does not send out any data to the shared bus, but uses a copy function of the memory to transport data from one partition to the other, using TCP/IP protocols.
- OptiConnect/400 is a more efficient way than traditional DDM using APPC controllers to access data on another AS/400 partition.

- At higher levels of throughput, the system overhead for OptiConnect/400 increases in a non-linear fashion. This is due to multiprocessing contention when CPU utilization is higher. For this reason, we recommend that CPU utilization should not exceed 90% on either the application machine or database server system.
- There is an overhead (called the *client/server factor*) associated with using OptiConnect/400 both on the application and database machine. This overhead varies and depends on such factors as:
  - Percent of database activity
  - Number of logical I/Os per transaction
  - Usage of journal and commitment control
  - Number of database opens and closes
- When analyzing batch workloads, pay more attention to logical database I/Os, blocking, and pool size. Performance characteristics may be different for batch versus interactive use of OptiConnect/400. In particular, batch jobs are often database intensive, and therefore, accumulate CPU and I/O delays for each logical I/O that is requested from the other system. This can add significant CPU utilization and elapsed time to batch runs.
- A detailed analysis to determine where to place the data files is critical in batch workloads or read-only files.
- In most distributed environments, including OptiConnect/400, there is a penalty for remote access. Obviously this penalty is most severe when 100% of accesses are remote and may become negligible for lower percentages (for example, 10%). Therefore, an optimum environment should minimize remote access by distributing users or data across systems in a manner to maximize local access. Carefully consider the replication of read-only files to application machines, the distribution of data across multiple database servers, and the location of batch work on application machines or database servers when defining an OptiConnect/400 environment to optimize total system capacity.

6.6 Available Tools

There are no tools to provide consolidated performance and capacity planning across a multi-partitioned system. Any result produced by such a tool would be meaningless. Logical partitions are independent systems and, as such, they must be managed as if they were separate, physical systems. It follows that all standard tools currently available to analyze system performance or performing capacity planning functions on physical machines can also be used most naturally on any single partition independently.

The same methodology applies. For each partition, you need to:

- Establish performance objectives
- Develop a good strategy to approach the problem
- Measure current performance
- Perform data analysis
- Perform trend analysis
- Perform problem analysis
- Setup benchmark for progress evaluation

In the next few sections, we explain some of the tools available and go into detail for those such as Management Central, which have been extended to facilitate the task of managing performance on multi-partitioned and clustered systems.

6.6.1 Performance Tools

The role of Performance Tools has not changed. In the initial planning stages to implement a logical partitioning configuration, Performance Tools are used for performance data collection, performance data analysis, and problem analysis. These steps are still done on each single partition in the same way as they are done on any non-partitioned system.

6.6.2 BEST/1

BEST/1 is the tool of choice to help determine hardware requirements as a preliminary to creating a logically partitionable machine.

After you work out the broad lines of your configuration, you may not be able to find a physical machine that matches your preferred configuration. For example, you may have decided that you need a 7XX model with a three-processor logical partition. This is not a valid physical machine, but you still need to measure how such a logical machine will handle the workload that you are going assign to it.

The solution to this dilemma is to create a new customized processor feature code using BEST/1. The resulting configuration can then be used for standard modelling. The next section explains the methodology to do it.

6.7 Creating a New Hardware Feature in BEST/1

If you created your own individual logical partition on an AS/400 model 7XX, you cannot find this model information in the hardware file in BEST/1. In BEST/1, you will find all of the 7xx models and related feature codes you can order, for example, an AS/400 Model 740 with a processor feature code of

2E6D. However, if you create a logical partition on a Model 720 with, for example, two processors that each have 400 CPW and your interactive performance will be at 46%, you have to create your own specific hardware feature in BEST/1. This section guides you through the setup of a new hardware feature in BEST/1.

6.7.1 Going to the Partition Processing Configuration

To go to the partition processing configuration, you have to start System Service Tools (SST). To do this, complete these steps:

1. Enter the command:

   ```
   STRSST
   ```

2. Select option **5**, Work with system partitions (Figure 38).

```
                    System Service Tools (SST)

Select one of the following:

     1. Start a service tool
     2. Work with active service tools
     3. Work with disk units
     4. Work with diskette data recovery
     5. Work with system partitions




Selection
     5

F3=Exit         F10=Command entry         F12=Cancel
```

Figure 38. Going to the Partition Processing Configuration (Part 1 of 4)

3. On the Work with System Partitions screen, select option **1** (Figure 39 on page 101).

```
                    Work with System Partitions
                                                              System:   LPAR2
Attention:  Incorrect use of this utility can cause damage
to data in this system.  See service documentation.

  Number of partitions  . . . . . :  2
  Partition manager release . . . . . :  V4R4M0

  Partition identifier  . . . . . :  1
  Partition name  . . . . . :  LPAR2

Select one of the following:

     1. Display partition information
     2. Work with partition status

     4. Recover configuration data


Selection
     1

F3=Exit   F12=Cancel
```

Figure 39. Going to the Partition Processing Configuration (Part 2 of 4)

4. On the Display Partition Information screen, select option **2** (Figure 40).

```
                    Display Partition Information
                                                              System:   LPAR2
Select one of the following:

     1. Display partition status
     2. Display partition processing configuration
     3. Display allocated I/O resources
     4. Display available I/O resources




Selection
     2

F3=Exit   F12=Cancel
```

Figure 40. Going to the Partition Processing Configuration (Part 3 of 4)

```
                Display Partition Processing Configuration
                                                          System:   LPAR2
Number of system processors . . . . . . . . . . . :  4
Size of system main storage (MB)  . . . . . . . . :  8192

Partition            Total       Main Storage  Interactive  Inter-partition
Identifier  Name     Processors  Size (MB)     Percentage   OptiConnect
  1         PAR2       2         4096           46            Yes
```

```
Only this partition is shown, use primary to see all partitions.
F3=Exit   F11=Display allocated I/O resources   F12=Cancel
```

Figure 41. Going to the Partition Processing Configuration (Part 4 of 4)

You now have all the information you need to create a new hardware feature in BEST/1.

In our example, we have a base system with four processors and 8,192 MB of main storage. Actually, this AS/400 system is a Model 720 with processor feature code 208D. In Figure 42 on page 104, you can see the relative performance rating of this model. The processor performance value is 403.69. The value for the interactive performance is 219.41.

The performance values in BEST/1 are still based on Model B10, which has a relative performance rating of 1.0. The published CPW values for all of the AS/400 systems cannot be found in BEST/1. We do not need them for our new hardware feature. We need the performance values found in Figure 42 for the calculation of the new hardware feature in BEST/1.

6.7.2 Calculating New Performance Ratings for Logical Partitions

To calculate the new performance ratings, follow the process described here:

The *processor performance* rating is 403.69. The base system has four processors. Determine the performance rating by using this equation:

```
403.69 / 4 = 100.92
```

This produces a performance rating of 100.9 for each processor.

Now calculate the relative performance rating for the batch:

```
100.92 * 2 = 201.84
```

This produces a relative performance rating for batch of 201.84. The *interactive performance* rating is 219.41. The interactive percentage in logical partition is 46%. 46% of 219.41 is 100.93.

The maximum available relative performance rating is 201.84. Our interactive performance rating is 100.93. These are the values that we need when we create our new hardware feature in BEST/1.

6.7.3 Creating the New Hardware Feature in BEST/1

To create a new hardware feature member in the BEST/1 hardware file, follow these steps:

1. Type `STRBEST` and press **Enter**.
2. Press **Enter** on the disclaimer screen.
3. On the BEST/1 for the AS/400 screen, type `60` (More BEST/1 Options) and press **Enter**.
4. On the More BEST/1 Options screen, type `10` (Hardware characteristics menu) and press **Enter**.
5. On the Hardware Characteristics screen, type `1` (Work with CPU models) and press **Enter**.
6. On the list of CPU models, find one with the same CPU value (in our example, 720) and a model value that comes close to the one noted (in our example, 208D).
7. On the Work with CPU Models screen, type `3` (copy) next to the model that you selected in the previous step (Figure 42 on page 104) and press **Enter**.

```
                              Work with CPU Models

Type option, press Enter.
  1=Create   2=Change   3=Copy   4=Delete   7=Rename

                  System   Archi-  --Relative Perf---   Number of  Max Stor
Opt  CPU Model    Unit   tecture    Normal    Server  Processors    (MB)

     720   207E   9406    *RISC     104.27    199.86       2         8192
     720   207F   9406    *RISC       9.05    403.69       4         8192
     720   208A   9406    *RISC      30.77    403.69       4         8192
     720   208B   9406    *RISC      61.55    403.69       4         8192
     720   208C   9406    *RISC     109.70    403.69       4         8192
 3   720   208D   9406    *RISC     219.41    403.69       4         8192
     436   2102   9402    *RISC       4.80                 1          224
     436   2104   9402    *RISC       6.10                 1          224
     436   2106   9402    *RISC       8.70                 1          256
     170   2159   9406    *RISC       3.60     22.50       1          832
     170   2160   9406    *RISC       5.10     29.30       1          832
     S20   2161   9406    *RISC       6.50     29.30       1         2048
     S20   2163   9406    *RISC       8.00     57.70       1         2048
                                                                      More...
F3=Exit   F9=Display equivalent BEST/1 and IBM CPU models   F12=Cancel
```

Figure 42. Work with CPU Models — BEST/1

8. On the Copy CPU Model screen, type in the new name (in our example, LPAR) and press **Enter**. You will see a *208D* is copied at the bottom of the screen (Figure 43).

```
                               Copy CPU Model

To copy, type New Name, press Enter.

Current name  . . . . . . . . . . :   208D

New name  . . . . . . . . . . . . .   LPAR









F3=Exit   F12=Cancel
BEST/1 CPU model 720 208D refers to IBM CPU model 720 2064-1505
```

Figure 43. Copy CPU Model — BEST/1

9. Edit the new model you created by typing a 2 next to it on the Work with CPU Models screen (Figure 44). Press **Enter**.

```
                              Change CPU Model

CPU model  . . . . . . . . . . . . . . . . :   LPAR
Min/Max storage size (MB)  . . . . . . . . :   256     8192

Type information, press Enter.
  System unit . . . . . . . . . . . . .   9406         9402, 9404, 9406
  Architecture  . . . . . . . . . . . .   *RISC        *CISC, *RISC
  Relative performance (B10 = 1.0):
    Normal  . . . . . . . . . . . . . .   100.93
    Server  . . . . . . . . . . . . . .   201.84       (Blank if not Server)
  Number of processors  . . . . . . . .   2
  Currently available . . . . . . . . .   Y            Y=Yes, N=No
  Family  . . . . . . . . . . . . . . .   *POWERAS     Name
  Upgrade to family . . . . . . . . . .   *NONE        *NONE, name


                                                    Minimum     Maximum
  Disk IOPs  . . . . . . . . . . . . . . . . . .    0           22
  Multifunction IOPs . . . . . . . . . . . . . .   1           17

F3=Exit   F6=Specify storage sizes       F9=Specify connections to disk IOPs
F11=Specify connections to disk drives   F12=Cancel   F24=More keys
BEST/1 CPU model 720 LPAR refers to IBM CPU model 720 2064-1505
```

Figure 44. Change CPU Model — BEST/1

10. On the Change CPU Model screen, update the performance values to the ones we calculated before. Change the number of processors to 2, and change the value Upgrade to family to *NONE. Then, press **Enter**.
11. From the Work with CPU Models screen, press **Enter**.
12. From the Hardware Characteristics screen, press **F15** (Save current hardware characteristics).
13. On the Save Hardware Characteristics screen (Figure 45 on page 106), enter a member name, (in our example, NEWLPAR). Leave the library name as QGPL. Press **Enter**.

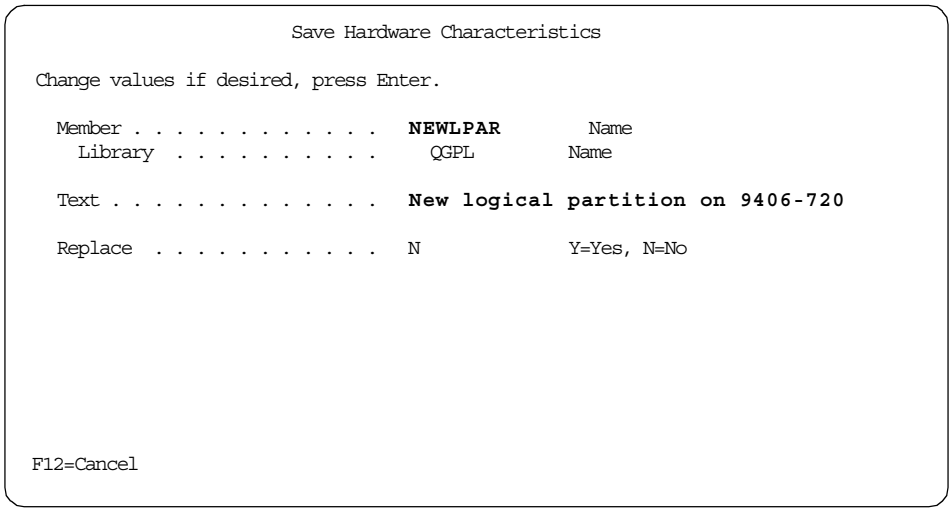


Figure 45. Save Hardware Characteristics — BEST/1

14. From the Hardware Characteristics screen, press **Enter**.
15. From the More BEST/1 Options screen, press **Enter**.
16. From the BEST/1 for AS/400 screen, press **F3**.

    You should be back at the command line.
17. Enter the following command:

    ```
    WRKMBRPDM QGPL/QACYHRWS
    ```

    The screen shown in Figure 46 on page 107 appears.

```
                        Work with Members Using PDM                RCHASM25

File  . . . . . .   QACYHRWS
  Library . . . .     QGPL                Position to  . . . . .

Type options, press Enter.
  3=Copy    4=Delete         5=Display   7=Rename    8=Display description
  9=Save    13=Change text   18=Change using DFU     25=Find string ...

Opt  Member      Date        Text
 3   NEWLPAR     05/04/99    New logical partition on 9406-720



                                                                       Bottom
Parameters or command
===>
F3=Exit          F4=Prompt            F5=Refresh            F6=Create
F9=Retrieve   F10=Command entry   F23=More options   F24=More keys
```

Figure 46. Work with Members Using PDM

18.On the Work with Members Using PDM screen, type a `3` next to the *NEWLPAR* member. Press **Enter**. The screen shown in Figure 47 appears.

```
                              Copy Members

From file . . . . . :   QACYHRWS
  From library  . . . . :     QGPL

Type the file name and library name to receive the copied members.

  To file . . . . . .   QACYHRWS     Name, F4 for list
    To library  . . . .     QPFR

To rename copied member, type New Name, press Enter.

Member         New Name
NEWLPAR        NEWLPAR



                                                                       Bottom
F3=Exit             F4=Prompt     F5=Refresh     F12=Cancel
F19=Submit to batch
```

Figure 47. Copy Members

19.On the Copy Members screen, type `QPFR` in the *To Library* parameter. Press **Enter**.

20. On the Work with Members Using PDM screen, press **Enter** to go to the command line.

Using these steps, you created a new hardware model in BEST/1. If you are now sizing a workload or modelling with performance data for the logical partition, you have the exact hardware characteristics of the logical partition in BEST/1.

6.7.4 Using the New Hardware Feature

If you started a performance collection using *Collection Services* function of Management Central and you created the database files for the Performance Tools on the AS/400, you can create a BEST/1 model from this performance data. This performance data will have all of the right information about the hardware of the partition on which you collected the data. The only values that are not included are the relative performance ratings for this partition. If you are going to use the collected data for analyzing and modeling using BEST/1, you have to go to the configuration menu. Using option 1, you can change your CPU and other resource values. Figure 48 shows how to specify the created hardware feature from the Change CPU and other Resource Values display in BEST/1.

```
                    Change CPU and Other Resource Values

Model/Text:   LPAR2       Test TCP/IP FTP on RCHASM25

Type choices, press Enter.

  System unit . . . . . .   9406      9402, 9404, 9406
  CPU Model . . . . . . .   LPAR       F4 for list
  System storage (MB) . .   4096      F4 for list
  Release level . . . . .   V4R4M0    F4 for list

F3=Exit   F4=Prompt   F12=Cancel
```

Figure 48. Change CPU and Other Resource Values

If you press **F4** and your previously created hardware feature is not shown in the available hardware features, go to the Hardware Characteristics menu and press **F14**. Your created model will appear in the list. Figure 49 shows the Hardware Characteristics menu.

```
                          Hardware Characteristics

 Select one of the following:

      1. Work with CPU models
      2. Work with Disk IOP features
      3. Work with Disk Drive features
      4. Work with Communications IOP features
      5. Work with Communications line speeds
      6. Work with Multifunction IOP features
      7. Work with IOA features









 Selection or command
 ===>

 F3=Exit   F4=Prompt   F9=Retrieve   F12=Cancel
F14=Replace with saved hardware characteristics  F24=More keys
```

Figure 49. Hardware Characteristics Menu

6.7.5 Performance Management/400

Performance Management/400 (PM/400) offers a broad array of benefits. It is a systems management tool that ensures that you get the most out of your system by regularly monitoring performance and growth. These measurements allow you, your IBM Business Partner, or IBM to more easily plan future system growth and identify potential resource constraints. Measuring your system's performance helps you make decisions that affect your budget and human resources plans.

PM/400 is automated, self-managing, and easy to use. PM/400 gathers the non-proprietary performance and capacity data from the AS/400 system and automatically sends it to IBM. In return, you receive this data in a series of reports and graphs that show your system's growth and performance.

PM/400 is already installed on your systems. It comes with the operating system, so it is installed in all your partitions. To send data, it uses the same resources as the Electronic Customer Support line. The IBM "factory" that

produces the reports is not set up to handle data from several partitions in this release, V4R4, since it is driven by a serial number.

Note: All partitions on a machine have the same serial number.

Only data from the primary partition is handled unless special arrangements are made with IBM. Since this is a restriction in the "factory" and not in the PM/400 code on your system, it may be subject to change. All features in PM/400 on your system may be used in all your partitions if it is enabled.

6.8 Performance Management Using Management Central

Management Central is an optionally installable, free component of Operations Navigator. Be sure to select Management Central when you install Operations Navigator. To do so, select one of the following methods to install Operations Navigator:

- **Full Install** — Installs Management Central
- **Custom Install** — Installs Management Central, but you have to select all subcomponents of Operations Navigator

If you already installed Operations Navigator using Typical Install or you did not select all components during Custom Install, you can always install Management Central later. Go to your Client Access Folder and open **Selective Setup**. Use the Selective Setup Wizard to install Management Central.

Once you install Management Central, it appears in the tree hierarchy in your Operations Navigator window. You need to expand Management Central to access it. For more information on how to use Management Central, refer to the Operations Navigator home page at: `http://www.as400.ibm.com/oper_nav`

6.8.1 Setting Up Your Central System

Management Central allows you to manage multiple AS/400 systems from a single AS/400 system in a TCP/IP network environment. This is the best way to manage a single AS/400 system with logical partitions. The logical partitions must be part of the TCP/IP network to be managed by Management Central. The primary partition should act as the central system to manage the logical partitions. The logical partitions are called *endpoint systems* in this kind of network.

To set up the primary partition as your central system, you must have authority to use the program that starts the Management Central server

(program QYPSSTRS in library QSYS). The first time you start Management Central, you must choose your central system.

To change your central system, right-click on **Management Central** and select **Change Central System**.

6.8.1.1 Adding Logical Partitions

A logical partition acts as an endpoint system. Endpoint systems are managed by the central system. When you first start Management Central, right-click **AS/400 Endpoint Systems** and select **Discover Systems**. This adds any connected AS/400 systems to your network as endpoint systems.

If you want to add logical partitions, manually as endpoint systems to your Management Central Network, right-click **AS/400 Endpoint systems**.Select **New Endpoint System**. Enter the system name of the specific logical partition.

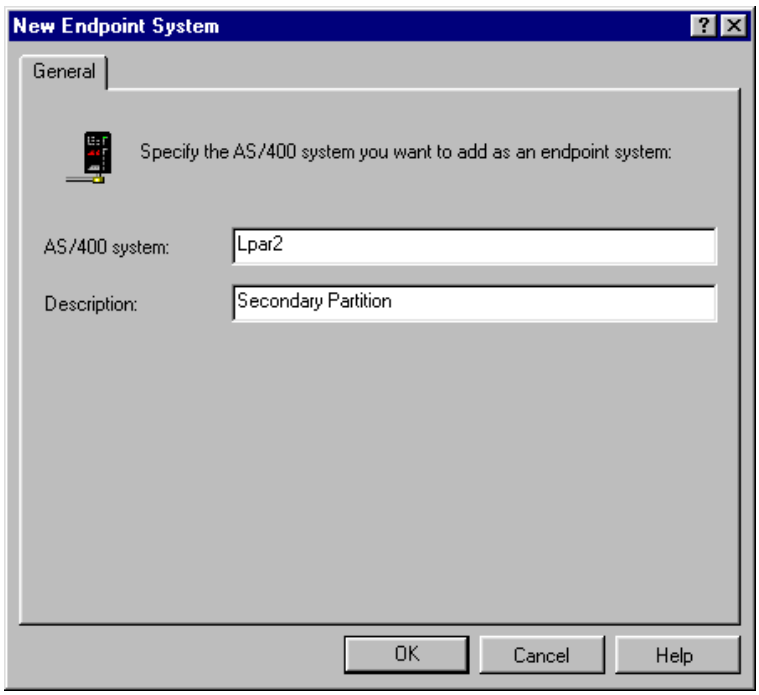


Figure 50. Adding an Endpoint System in Management Central

6.8.1.2 Creating System Groups

You can define a system group which is a collection of endpoint systems. Endpoint systems can belong to several system groups at once. The

advantage of the system group is that it can be managed from the central system just like a single system. If you add your logical partitions on your AS/400 system to a system group, this group can be easily managed by your primary partition.

To create a system group, complete these steps:

1. In the Operations Navigator window, open Management Central.
2. Right-click on **AS/400 System Groups** and select **New System Group**.
3. On the New System Group dialog, specify a name for the new system group. The name of the system group must be unique.
4. Select the logical partitions from the endpoint systems listed in the Available systems list. Click the **Add** button to add the logical partitions to the Selected systems list (see Figure 51).
5. Click **OK** to create the new system group with your logical partitions.

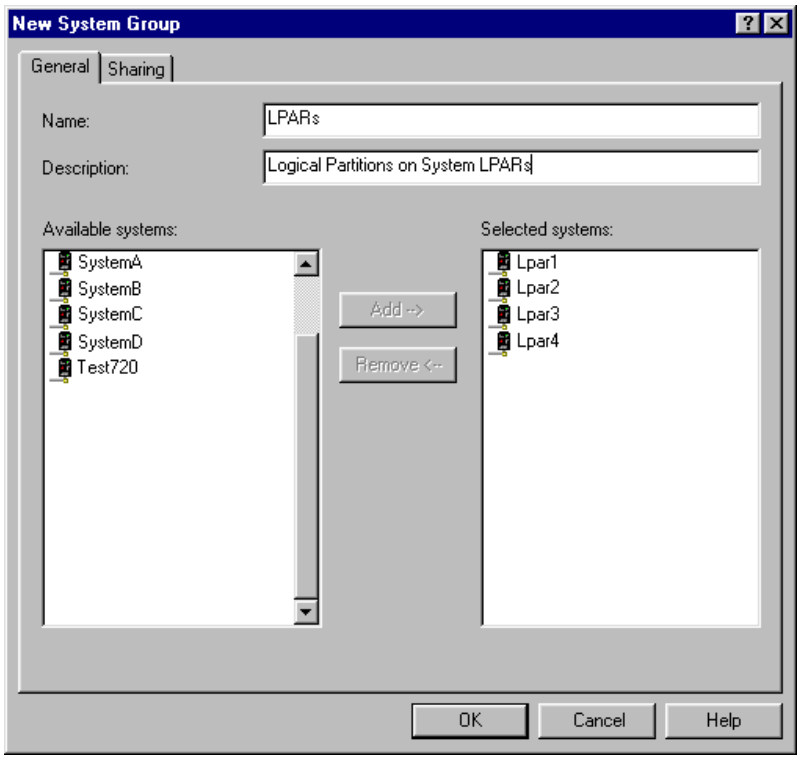


Figure 51. Creating New System Group

6.8.2 Collecting Performance Data

To collect performance data, use Management Central Collection Services. Collected data can be used by the Performance Tools for AS/400 licensed program product (5769-PT1). If you prefer realtime viewing of your performance data, Management Central provides an easy-to-use graphical interface for monitoring systems performance.

You can start Collection Services on a single system or you can start Collection Services on system groups. Collection Services collects information about the operation of your system that can be used to understand response time and throughput. This performance data can be used to make adjustments to programs and operations. It can also help you improve response times and throughput.

Performance data can be collected by using the OS/400 Start Performance Monitor (STRPFRMON) command. Using this command, 30 database files are created to collect performance data. Collection Services stores your data for each collection in a single collection object from which you can create as many different sets of database files as you need. You can use the database files with the Performance Tools for AS/400 licensed program (5769-PT1).

Collection Services also lets you:

- Manage your collection objects with the Operations Navigator graphical interface.
- Collect performance data continuously.
- Control what performance data is collected and how it is used.
- Move performance data between releases without converting the data.

6.8.2.1 Starting Collection Services on a Single Partition

To start collecting system performance data, you have to start Collection Services. Follow these steps to start Collection Services on a single partition:

1. Start Management Central in the AS/400 Operations Navigator window.
2. Expand Endpoint Systems.
3. Expand the partition where you want to collect the performance data.
4. Expand Configuration and Services.
5. Right-click **Collection Services** and select **Start Collecting** (Figure 52 on page 114).

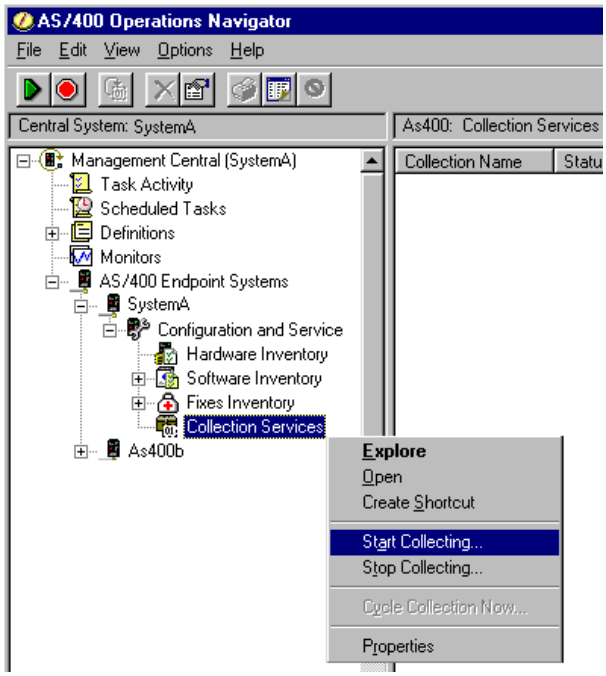


Figure 52. Start Collecting Performance Data on an Endpoint System

6. The default retention period for collection objects is one day. You may want to specify `Permanent` on the General Page, if you do not want Collection Services to delete your objects for you.
7. Click the **Data to Collect** tab.
8. Collection Services provides profiles to use for you. Use the Standard plus profile to collect all the data categories typically needed by the tools in Performance Tools for AS/400 (5769-PT1). To collect a smaller set of categories, select **Standard** or **Minimum**.
9. Click **OK** to start Collection Services immediately or click **Schedule** to start Collection Services at a later time.

All the performance data collected by Collection Services is stored in a single collection object. You manage your collection objects in the Operations Navigator graphical interface.

6.8.2.2 Starting Collection Services on a System Group

If you want to collect performance data on more than one partition, you can start Collection Services on a system group. The system group will contain all the logical partitions from which you want to collect performance data.

Follow these steps to start Collection Service on a system group containing logical partitions:

1. Start Management Central in the AS/400 Operations Navigator window.
2. Expand AS/400 System Groups.
3. Expand the partition where you want to collect the performance data.
4. Expand Configuration and Services.
5. Right-click **Collection Services** and select **Start Collecting**.
6. The default retention period for collection objects is one day. You may want to specify Permanent on the General Page, if you do not want Collection Services to delete your objects for you.
7. Click the **Data to Collect** tab.
8. Collection Services provides profiles for you to use. Use the Standard plus protocol to collect all of the data categories typically needed by the tools in Performance Tools for AS/400 (5769-PT1). To collect a smaller set of categories, select **Standard** or **Minimum**.
9. Click **OK** to start Collection Services immediately, or click **Schedule** to start Collection Services at a later time.

6.8.2.3 Customizing Data Collections

You can control what kind of data is collected by Collection Services and how often it is collected. You can select from the standard profiles provided by Collection Services. The profiles correspond to the settings in the OS/400 Start Performance Monitor (STRPFRMON) command for all data or for system data.

You can select **Custom** to create your own customized profile from a list of available data categories, such as System CPU, Local Response Time, Disk Storage, and IOPs. You can specify how often data is collected for each category of data.

For each category of data that you collect, you can specify how often the data will be collected. For many categories, you want to select the default collection interval, which you can set from predefined settings between 15 seconds and 60 minutes. The recommended setting is 15 minutes (see Figure 53 on page 116).

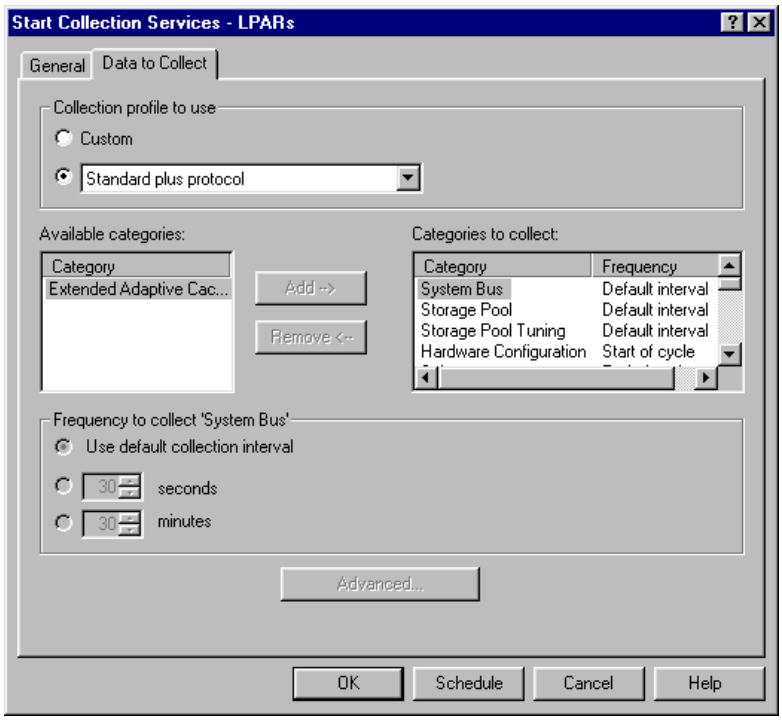


Figure 53. Start Collection Services on a System Group

6.8.2.4 Creating Database Files

Database files can be created automatically using Collection Services to collect performance data. You can also create database files from the collection object, where the data is stored after it is collected.

If you want to create your database files automatically as data is collected, select **Create database files** on the Start Collection Services dialog.

To export performance data from a collection object to database files, perform the following steps:

1. In AS/400 Operations Navigator, expand an AS/400 system. You can select an AS/400 endpoint system under Management Central or an AS/400 system to which you have a direct connection under My AS/400 Connections (or your active environment). Select the logical partition where you want the database file to be created.
2. Expand Configuration and Service.
3. Click **Collection Services**.

4. Right-click the collection object you want to export to database files, and select **Create Database Files**.
5. On the Create Database Files dialog (Figure 54), select the categories from the collection object to include in the database files. You can also select a different time period and sampling interval, as long as the collection object contains data to support your selections.
6. Click **OK**.

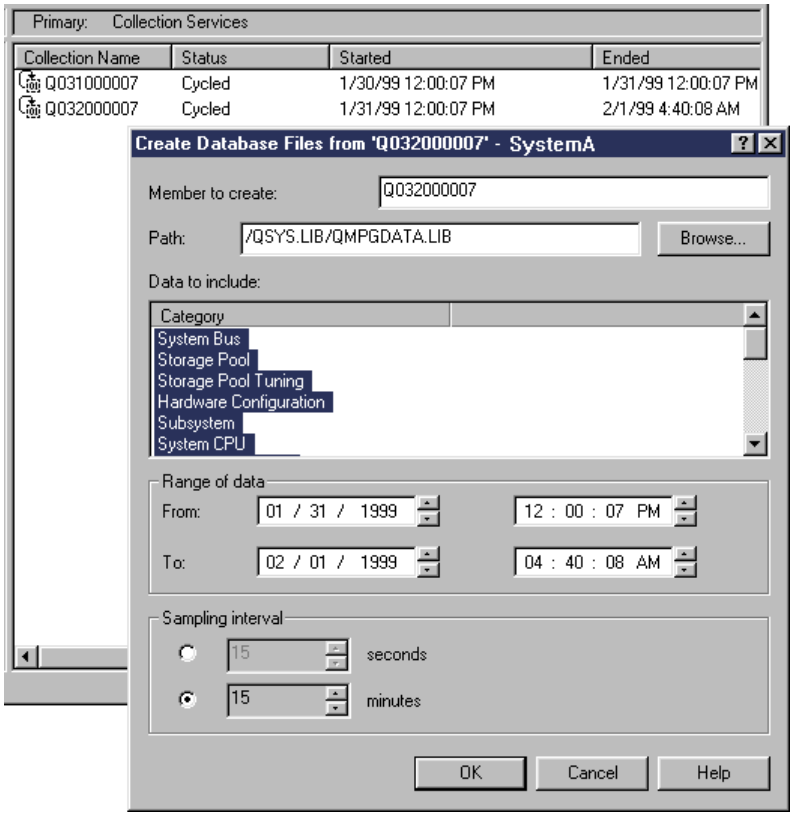


Figure 54. Creating Database Files from Collection Services

The job to create the database files is submitted to that logical partition where the performance data was collected. If you collected performance data for a System Group, there will be a collection object for each system or partition in this group. You have to create database files for each system or for each partition separately. There are no additional tools within Management Central for analyzing these database files.

6.9 Monitoring Realtime System Performance

Management Central provides a realtime performance monitor for AS/400 systems. This tool can be easily used to present realtime performance data on logical partitions. To collect performance data for later analysis, use Collection Services. The realtime performance monitor shows the performance data as it happens. You cannot save this information for later analysis.

6.9.1 Management Central

Management Central monitor graphs present system performance data in an easy-to-use graphical interface that you can directly manipulate to get different or more detailed data. Monitors allow you to collect performance data simultaneously for a wide variety of system metrics, for any system or system group, and for any length of time. Once you start a monitor, you are free to do other tasks on your AS/400 system, in Operations Navigator, or on your PC. In fact, you can even turn your PC off! It will continue to monitor your systems and perform any threshold commands or actions you specified. Your monitor will run until you decide to stop it. To effectively monitor realtime AS/400 system performance, create a Management Central monitor.

The metrics shown in Table 15 can be selected for your monitor graphs.

Table 15. Metrics for Realtime Performance Monitor Graphs

Metric Name	Metric Description
CPU Utilization	This metric shows the percentage of available processing unit time consumed by jobs on your system.
Interactive Response Time	This metric shows the response time that interactive jobs experience on your system.
Transaction Rate	The average number of transactions per second completed by jobs on your system.
Batch Logical Database I/O	This metric shows the average number of logical database input/output (I/O) operations currently performed by batch jobs on the system.
Disk Arm Utilization	This metric shows the percentage of disk arm capacity currently used on your system during the time you collect the data.
Disk Storage	This metric shows the percentage of disk arm storage that is full on your system during the time you collect the data.

Metric Name	Metric Description
Disk IOP Utilization	This metric shows how busy the disk input/output processors (IOPs) are on your system during the time you collect the data.
Communications IOP Utilization	This metric shows how busy the communications input/output processors (IOPs) are on your system during the time you collect the data.
Communications Line Utilization	This metric shows the amount of data that was actually sent and received on all your system communication lines.
LAN Utilization	This metric shows the amount of data that was actually sent and received on all your local area network (LAN) communication lines.
Machine Pool Faults	This metric shows the number of faults per second occurring in the system's machine pool.
User Pool Faults	This metric shows the number of faults per second occurring in all of the user pools on the system.

6.9.1.1 Creating a New Monitor

To create a new monitor, complete these steps:

1. Expand Management Central, right-click on **Monitors** and select **New Monitors**.
2. Select **General** page to:
   a. Enter a unique name for your monitor. You can enter any name for your monitor, up to 256 characters long.
   b. Enter a description for your monitor. Giving your monitor a description may help you distinguish any particular monitor from the other monitors you have running. This can be especially useful when you have many monitors running at once or when you want to copy monitors from one logical partition to another.
   c. Choose your metrics. You can select any metric, a group of metrics, or all the metrics from the list to be included in your monitor. Your options are shown in Table 15 on page 118.

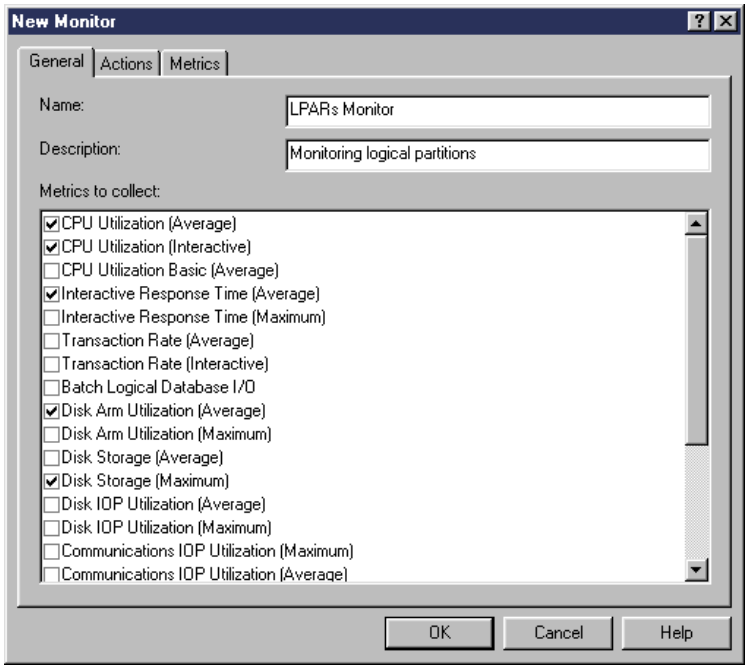


Figure 55. Adding a New Monitor

You can choose to run a monitor that contains any combination of these metrics on any AS/400 system, on any logical partition, or on system groups. After you name your monitor and select metrics, you are ready to view and change detailed metric information for each metric you selected for your monitor.

When you create a new monitor, you can specify actions that you want to occur on your PC when a threshold is triggered or reset. A threshold is a setting for a metric that is collected by a monitor. Threshold actions occur on your PC to notify you when threshold events occur. For example, you can choose to open monitor graphs automatically when thresholds are triggered. Threshold actions are different from any threshold commands that you may have set. Threshold commands run on your host system or client systems, while threshold actions occur on your PC.

When you are done working with the metric information on the New Monitor-Metrics page, be sure to click the **Edit Thresholds** button. This allows you to enable your thresholds and set threshold commands to run on the host or client when thresholds are triggered or reset. You can also use the

New Monitor-Actions page to set threshold actions to occur automatically on your PC when thresholds are triggered or reset.

To edit metric information, simply select the metric that you wish to edit from the drop-down list on the New Monitor-Metrics page *(*see Figure 56).

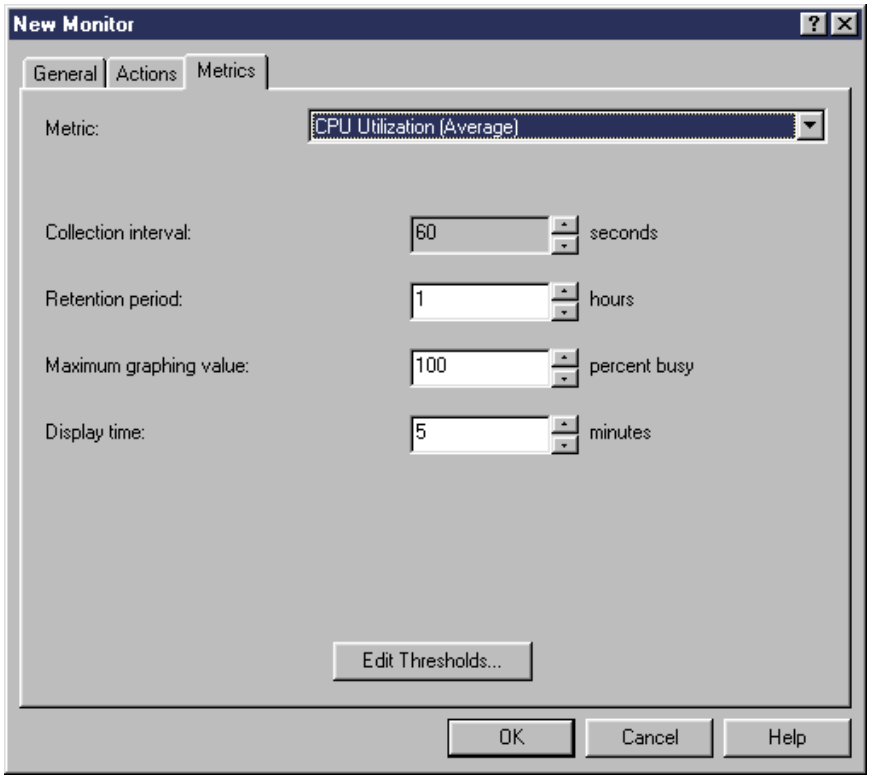


Figure 56. Edit Metric

Change any of the following properties:

- **Collection interval**

  This information specifies how often metric information is collected. Increase or decrease this number as needed.

- **Retention period**

  This information specifies how long the metric information will be stored on the host system. At the end of this period, the metric information will be deleted from the database even if your monitor is still running.

- **Maximum graphing value**

  This information specifies the highest value that will appear on the vertical axis of the monitor graph for this metric.

- **Display time**

  This information specifies the number of minutes that will appear on the horizontal axis of the monitor graph for this metric.

To edit the tresholds for a selected metric, click the **Edit Threshold** button. Use the threshold window to enable your thresholds. Before you can set any threshold commands, you must turn on your thresholds by selecting the **Enable Threshold** option. You can then use this window to enter any commands you want to occur when threshold trigger and reset values are reached.

6.9.1.2 Threshold Trigger Options

Threshold actions allow you to determine what you want to happen on your PC when your metric reaches a certain value (called the *trigger value*). You can also specify what you want to happen on your PC when your metric reaches a second value (called the *reset value*). You can choose to add events to the Event Log when your metrics reach trigger or reset values. You can set alarms on your PC. You can even set monitor graphs to open automatically on your PC when your metrics reach trigger or reset threshold values. Figure 57 on page 123 shows an example of the New Monitors-Actions tab.

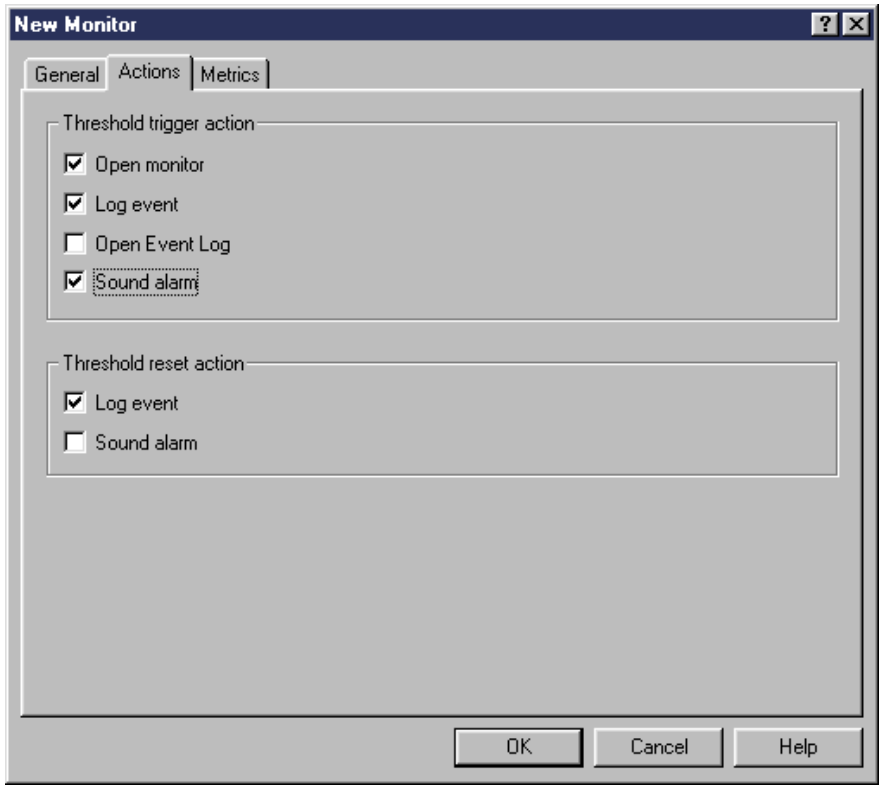


Figure 57. Treshold Trigger Options

Table 16 on page 124 provides an explanation for each action.

Table 16. Threshold Trigger Options

Action	Result
Open monitor	Displays the monitor graph when system performance reaches a threshold trigger for a particular metric. This allows you to see a graphical view of your system performance data as it is being collected. You do not have to keep the monitor graph open on your PC all the time. It opens automatically if you select this action, and you can keep the monitor graph open even if you close Operations Navigator.
Log event	Adds an entry to the Event Log on the central system indicating that the threshold was triggered.
Open Event Log	Displays the Event Log on your PC when a threshold trigger occurs. Much like the Open Monitor function, this action opens the Event Log only when you really need it.
Sound alarm	Sounds an alarm on your PC when system performance reaches a trigger value.

6.9.1.3 Threshold Trigger and Threshold Reset

The Edit Thresholds window (Figure 58) provides a place for you to specify the conditions that must be met to trigger and reset this threshold and the commands that should be run when the threshold is triggered and when it is reset. You cannot specify or change these conditions and commands unless the Enable threshold box is checked. You cannot specify the threshold reset conditions and commands until you have specified Value and Duration for the Threshold trigger.

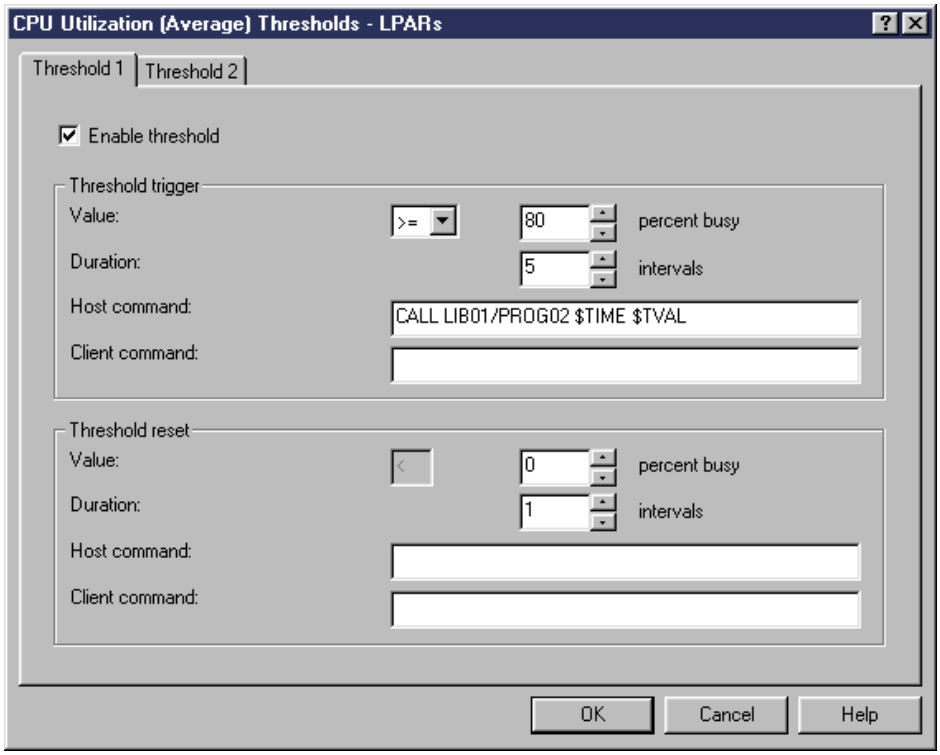


Figure 58. Edit Thresholds

You can specify the following conditions and commands for Threshold trigger and for Threshold reset:

- **Value**

  Specifies the condition that must be met to trigger or to reset this threshold.

- **Duration**

  Specifies the number of consecutive collection intervals that the value must meet the criteria to cause a threshold trigger or reset event. Specifying a higher number of collection intervals for Duration helps to avoid unnecessary threshold activity due to frequent spiking of values.

- **Host command**

  Specifies the command to be run on the AS/400 host system when the threshold is triggered or reset.

- **Client command**

  Specifies the command to be run on the PC when the threshold is triggered or reset. This command is run only if you are viewing this monitor on a PC when the threshold is triggered or reset. The command is run on every PC where you are running this monitor.

6.9.1.4 Threshold Commands

Use threshold settings to automate any command that you want to run when thresholds are triggered or reset. For example, you can set an AS/400 command that stops any new job from starting when CPU utilization reaches 90%. You can then set another command that allows new jobs to start when CPU utilization falls back to 70%. In another situation, you may have a monitor that collects data on average CPU utilization for a particular system. You can set thresholds and specify commands to keep the average CPU utilization between 20% and 90% or any boundaries you choose.

6.9.1.5 Parameters on Host or Client Command

You can use the following parameters with host commands or client commands to be run when a threshold is triggered or reset:

Parameter	Passed Data
$DATE	Date
$MON	Monitor name
$RDUR	Reset duration
$RVAL	Reset value
$INTVL	Collection interval
$SEQ	Sequence number
$TIME	Time
$TDUR	Trigger duration
$TVAL	Trigger value
$VAL	Current value

Consider the following example. The following host command uses the $TIME and $TVAL parameters to pass to the program the time that the threshold was triggered and the trigger value (see Figure 59 on page 127):

```
CALL LIB01/PROG02 ($TIME $TVAL)
```

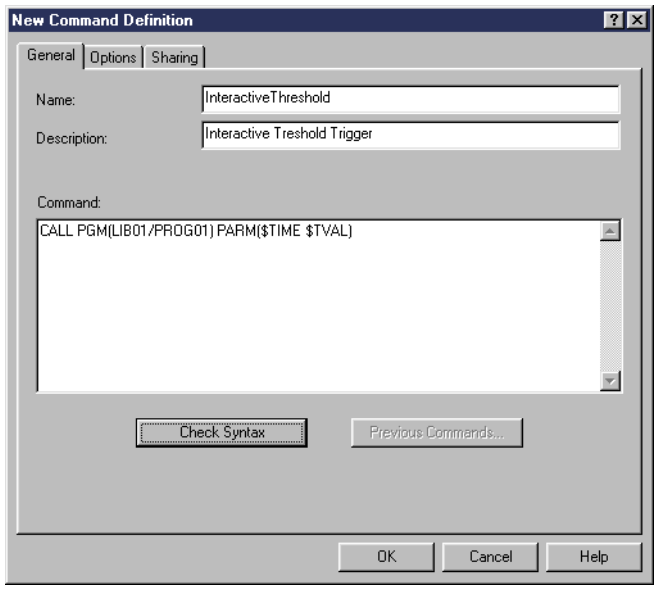


Figure 59. New Command Definition for Thresholds

6.9.1.6 Creating a Command Definition

Management Central enables you to define a command and then run it on multiple systems. To do this, you must first create your command definition, using a few quick steps:

1. In AS/400 Operations Navigator, expand Management Central.
2. Expand Definitions.
3. Right-click on **Command** and select **New Definition**.
4. Specify a name for the definition, a brief description, and the command to be run. You can click **Previous Commands** to select from a list of commands you previously ran, or you can click **Check Syntax** to validate the syntax of the command on the central system.
5. To specify options concerning the job log or inquiry messages, click the **Options** tab.
6. Click the **Sharing** tab to specify whether you want to share this package definition with other users.
7. Click **OK**.

After you create your command definition, you can run the command on systems or groups.

6.9.2 Running a Monitor

Once you create a monitor by selecting metrics and setting threshold actions, the next step is to run the monitor on your AS/400 systems. To start or stop a monitor in Management Central, select the monitor that you wish to start and use the start and stop icons in your toolbar. You can also follow the simple steps described here.

To start a monitor, follow these steps:

1. In Management Central, select **Monitors**.
2. Right-click the monitor you want to start, and select **Start**.
3. Select the systems or system groups on which you want to run the monitor, and click **OK**.

To stop a monitor, complete these steps:

1. In Management Central, select **Monitors**.
2. Right-click on the monitor that you want, and select **Stop**.

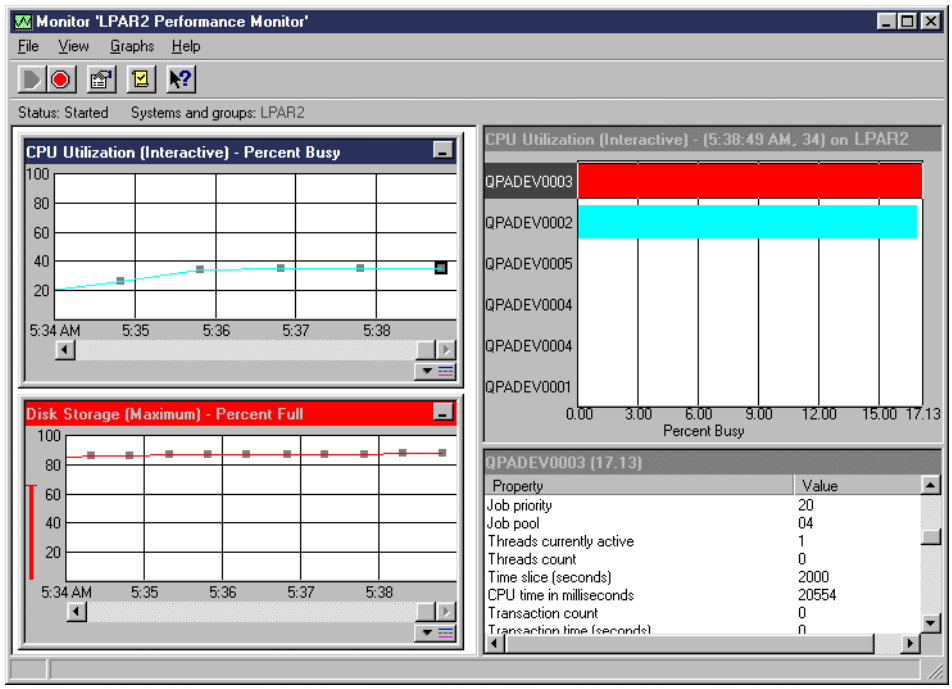


Figure 60. Performance Monitor — CPU Utilization and Disk Storage

When your monitor is up and running, work with the monitor graphs to see realtime system performance data in an easy-to-use graphical interface (Figure 60 on page 128).

You can use monitor graphs to find more information about system performance. For example, move your mouse over a collection point (or data point) within a graph to get information about the exact time and the AS/400 system for which the data was collected. When you click a data collection point in a monitor graph, highly detailed information for that data point appears in the upper right Details panel of the monitor window. The lower right Properties panel shows properties for the information that you select in the Details panel.

6.9.2.1 Advanced Tips for Monitor Graphs

Once you create a new monitor, you can pick up some tips to make monitoring your system or your partitions even easier. Management Central provides a great deal of flexibility and functionality that you can use to get the most out of your system performance data. For example, you can:

- Create shortcuts to your monitors
- Use the Event Log
- Create different graph line colors for each logical partition
- Change threshold values directly on your graph
- Sort the bars on your monitor graphs

The example in Figure 60 on page 128 shows you the Performance Monitor from one system. You can easily expand this view to show more than one system. This gives you the flexibility to watch and manage the performance on your system with logical partitions. For example, if you want to control the average CPU Utilization of your logical partitions, you simply have to create a new monitor with this metric and during start you select all your partitions or, even better, your system group with all the logical partitions. If you want to be notified if the CPU Utilization on a logical partition is more than a specific percentage, set a threshold, and the monitor of that partition becomes red if the threshold is reached. In Figure 61 on page 130, the Performance Monitor shows two logical partitions. Each partition is represented by its own line chart. Move the mouse over the line to get detailed information about the metrics.

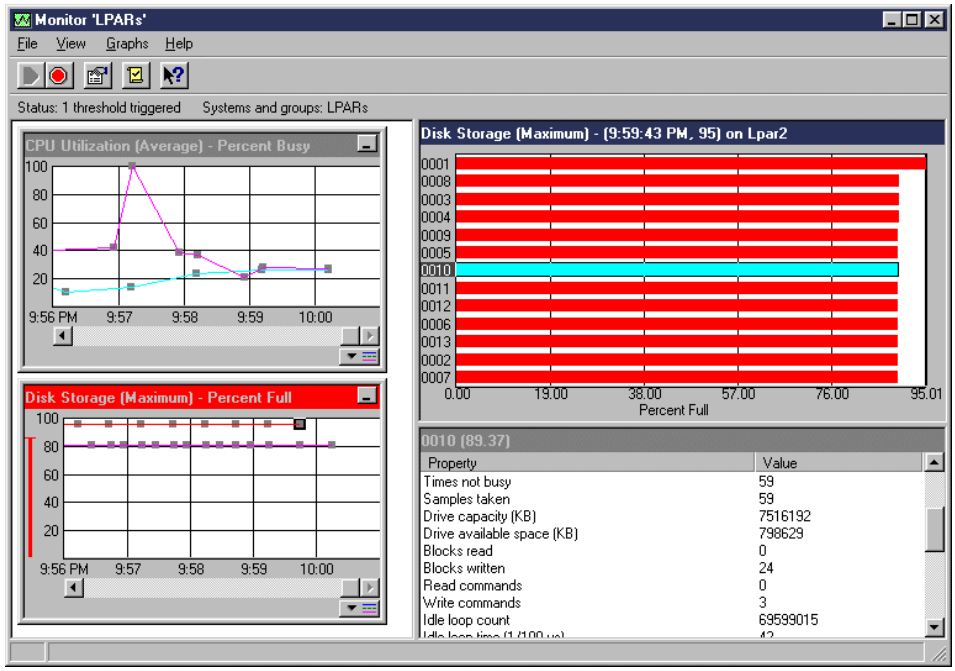


Figure 61. Monitoring Two Logical Partitions

6.10 On-Going Performance Monitoring

After the initial installation of a multi-partitioned system, close monitoring of the performance of each partition becomes a normal task, as with independent physical systems. However, a new dimension is introduced with logical partitioning.

The logically partitioned machine allows the movement of resources between partitions. During your on-going analysis, you may find that the workload in a three-processor partition has decreased to the point where two processors can do the work. You may now decide to reduce the resources in that partition and allocate them to another partition where the workload has increased. Once the changes are implemented, you need to analyze again the performance of both the giving and the recipient partitions to make sure that you achieve your objective. Logical partitioning does not alter this standard cycle to any great extent.

Another situation is where the workload on some partitions have increased, but there is no scope for re-allocating any existing resources. In this case, you would need to upgrade your physical machine and acquire additional

resources such as processors, main storage, and interactive performance, to cope with the workload. Incorporating additional resources in the existing configuration would be done using the standard technique of resource allocation to individual logical partition described elsewhere.

As we saw in 6.6, “Available Tools” on page 98, there is a rich array of tools available for systems management. Each site needs to develop its own performance management strategy. On a non-partitioned system, this process is applied to the entire physical machine. With a partitioned machine, it is repeated as many times as there are partitions on the physical machine.

Chapter 7. Inter-Partition Communication

This chapter explains how to implement virtual OptiConnect on primary and on secondary partitions. It explains the difference between implementing OptiConnect/400 with external buses and implementing OptiConnect/400 on virtual hardware. It also describes the capabilities of OptiConnect/400.

Inter-partition communication is not limited to virtual OptiConnect. You can use existing communications adapters in each logical partition to communicate to another logical partition. If you have Token-Ring or Ethernet adapters in your logical partitions, you can also use all the applications described in this chapter. However, to have the full advantage of high-speed communications between partitions, you should use virtual OptiConnect.

7.1 OptiConnect/400 Capabilities

The fastest way to communicate between logical partitions is to use virtual OptiConnect. To understand the methology of OptiConnect, there will be a brief introduction to using OptiConnect/400. The only difference between virtual OptiConnect and OptiConnect/400 is that virtual OptiConnect uses no additional hardware. Virtual OptiConnect is communication within one physical system divided into logical partitions. Virtual OptiConnect requires a single license for OptiConnect/400 that can be used by all partitions.

7.1.1 OptiConnect/400

OptiConnect/400 (see Figure 62 on page 134) is a combination of hardware and software that allows users to connect multiple high-end AS/400 systems using a high-speed fiber optic bus connection. OptiConnect hardware consists of OptiConnect Receiver cards that are installed in a dedicated I/O Expansion Tower. They are connected to Satellite Systems using fiber-optic cables.

OptiConnect software consists of the following items:

- Additions to OS/400 that provide fast path Distributed Data Management (DDM) access across an optical bus. OptiConnect/400 is a priced feature and must be installed as 5769SS1 (product option 23).
- APIs, if the OptiMover PRPQ (5799-FWQ) has been purchased, that can be coded to by third parties to use the hardware.
- A connection manager that manages OptiConnect resources.
- An agent job that runs on the server on behalf of client requests.

Two features differentiate OptiConnect from traditional communications-based distributed operations:

- The first is a system bus connection between multiple systems using high-speed fiber optic technology.
- The second is an I/O driver embedded in the operating system that streamlines the application access to data on a remote system. To accomplish this, OptiConnect provides a shared bus on which systems communicate using a peer-to-peer protocol. After OptiConnect establishes system connections on the shared bus, much of the advanced peer-to-peer communication (APPC) protocol stack is bypassed.

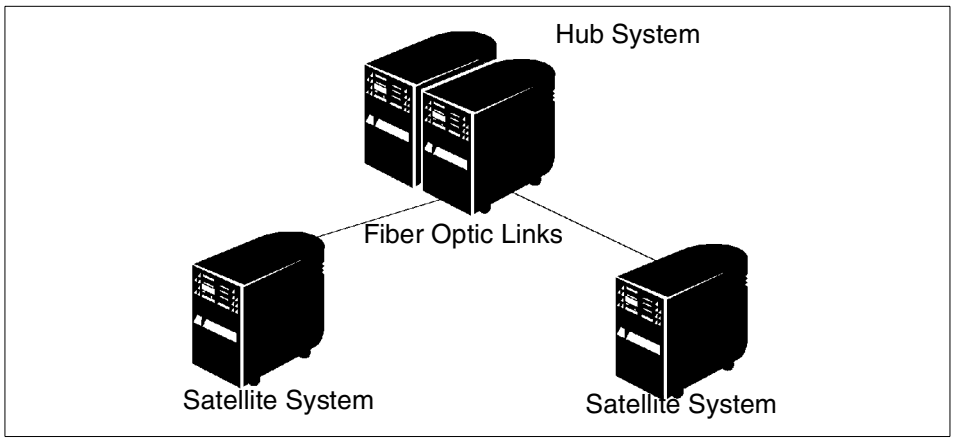


Figure 62. OptiConnect/400 Concept

7.1.1.1 OptiConnect/400 Bandwidth

Traditional communication protocol overhead is too impractical for heavy workload distributed applications. To minimize latency, or the time it takes to send a message and receive a response, the protocol chosen must be efficient in the number of code steps required. To maximize bandwidth, or the amount of data that can be transferred in a unit of time, high-performance hardware must be used.

The OptiConnect channel is very efficient and the best solution to both latency and bandwidth. The length of the cable affects how low latency can get. As distance increases, the speed of light becomes a limiting factor for a given bandwidth. OptiConnect is limited to 500 meters over the 1063 Mbps (megabit per second) link used on RISC AS/400 models. It is limited to two kilometers over the 220 Mbps link supported on CISC AS/400 models.

7.1.2 OptiMover for OS/400

OptiMover for OS/400 is a special version of OptiConnect for OS/400. It provides application programs with Application Program Interfaces (APIs) to the services of the OptiConnect Connection Manager and Device Driver. With the OptiMover for OS/400 PRPQ, applications may use the high-speed optical bus to communicate among systems on the shared bus, including the object transfer functions in ObjectConnect, but the high-speed DDM capabilities of OptiConnect are not available. Because OptiMover is essentially the same as the OS/400 OptiConnect feature, it uses the same combination of hardware and software that allows you to connect multiple AS/400 systems on a shared optical bus.

Note: The PRPQ number for OptiMover is P84291, program 5799-FWQ.

7.2 OS/400 Functions Using OptiConnect/400

The AS/400 has many functions available in its operating system that can use OptiConnect/400. The most common are Distributed Data Management/400 (DDM) and ObjectConnect/400 for concurrent save and restore of objects without using a tape media.

7.2.1 Distributed Data Management (DDM)

DDM is part of the OS/400. The DDM support of the AS/400 system allows application programs or users to access data files that reside on remote systems, and starting with OS/400 V4R4, also on logical partitions.

Distributed Data Management provides horizontal growth using OptiConnect/400 as a simple method of accessing files on other systems through the use of DDM files. There is a large overhead on CPU and response times using traditional DDM with common communication methods. IBM offers DDM at system bus speed by using OptiConnect/400.

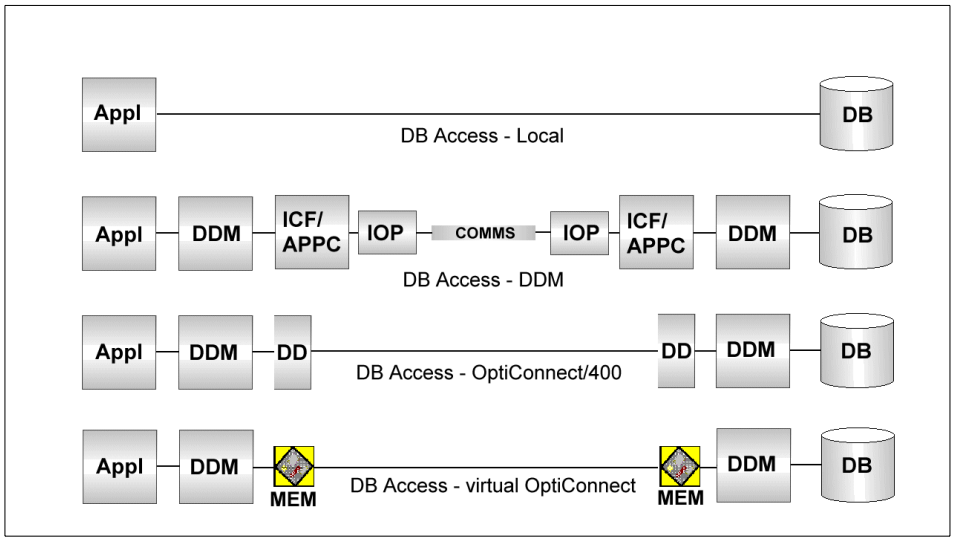


Figure 63. Distributed Data Management Concept

Figure 63 shows that there is an overhead using Distributed Data Management over a communication line. OptiConnect/400 uses a fast Device Driver (DD) and no ICF/APPC function is involved in a DDM transaction.

There are some limitations when implementing DDM of which you should be aware. Customers should read and understand these limitations which can be found in *Distributed Data Management*, SC41-3307.

Some of the most common commands are shown in the following list along with the details on how DDM affects them:

- **ALCOBJ**

  When using the Allocate Object (ALCOBJ) command over DDM, the object on the source system is locked as well as on the target side. When DDM-related files are being allocated, a longer time is required to complete the command.

- **CHKOBJ**

  Using the Check Object (CHKOBJ) CL command to check the existence of members across DDM does not work.

- **Multimembered files**

  DDM files do not have the option to access multiple members. You can do this by specifying the member name in the Override with Database File (OVRDBF) command before accessing the file.

- **Query/400**

  You can run an existing query across DDM. However, no changes can be made to it. The product Query/400 should not be used. If you need to run queries against DDM files, perform a Passthrough or Telnet to the other system. Then, run the query with all its options from there.

- **Remote data areas and remote data queues**

  Remote data areas and remote data queues are supported by OptiConnect/400. They can be used, but there must be an existing communications path from one partition to the other.

- **SBMRMTCMD**

  You can use the Submit Remote (SBMRMTCMD) command, but no attributes are passed over DDM.

- **SQL**

  SQL splits its workload between the two partitions. There are no problems using SQL over DDM. If you plan to use SQL across DDM, perform a benchmark first to predict the performance.

7.2.2 ObjectConnect/400

ObjectConnect/400 provides six specialized save and restore commands. With these commands, you get the ability to save and restore objects, libraries, DLOs, and IFS directories to a second partition. ObjectConnect/400 operates over several communication methods. Virtual OptiConnect always provides the best performance out of all of them. Some of the most common commands are described in the following list:

- **SAVRST**

  You can use the Save/Restore (SAVRST) command to save one or more objects in directories, send them to another system, and restore them. It can also save entire directories (not to be confused with entire systems). The SAVRST command supports the same options as the SAV command.

- **SAVRSTOBJ**

  You can use the Save/Restore Object (SAVRSTOBJ) command to save one or more objects, send them to another system, and restore them. The SAVRSTOBJ command supports the same options as the SAVOBJ command, including the use of the OMITOBJ parameter.

- **SAVRSTCHG**

  You can use the Save/Restore Changed Objects (SAVRSTCHG) command to save one or more changed objects, send them to another

system, and restore them. An example of this would be a situation where you want to maintain duplicate sets of files on two different systems. The SAVRSTCHG command supports most of the same options as the SAVCHGOBJ command, including the use of OMITOBJ parameters. You may use the OMITLIB parameter with this command. You may also specify generic values for the LIB parameter on this command.

- **SAVRSTLIB**

  You can use the Save/Restore Library (SAVRSTLIB) command to save one or more libraries, send them to another system, and restore them. The SAVRSTLIB command supports the same options as the SAVLIB command, including the use of the OMITLIB and OMITOBJ parameters. You may also use generic values for the *LIB parameter on this command.

- **SAVRSTDLO**

  You can use the Save/Restore Document Library Object (SAVRSTDLO) command to save one or more document library objects, send them to another system, and restore them. The SAVRSTDLO command supports the same options as the SAVDLO command.

- **SAVRSTCFG**

  You can use the Save/Restore Configuration (SAVRSTCFG) command to save one or more configuration objects, send them to another system, and restore them. The SAVRSTCFG command supports most of the options and parameters as the SAVCFG and RSTCFG commands do.

  When you copy your configuration by using the SAVRSTCFG command, the system saves and restores the following object types:

  – *CFGL Configuration List
  – *COSD Class-of-Service Description
  – *CNNL Connection List
  – *MODD Mode Description
  – *NTBD NetBIOS Description
  – *IPXD IPX Description

7.2.3 Remote Journal Function

The remote journal function allows you to establish journals and journal receivers on a remote AS/400 system that are associated with specific journals and journal receivers on a local system. The remote journal function can replicate journal entries from the local system to the journals and journal receivers that are located on the remote system after they are established.

The remote journal function helps to efficiently replicate journal entries to one or more systems. You can use the remote journal function with application programs to maintain a replica database. A replica database is a copy of an original database that resides on another AS/400 system. An original database resides on a primary system. Applications make changes to the original database during normal operations. Previous to V4R2M0, you could accomplish a similar function by using the Receive Journal Entry (RCVJRNE) command. The remote journal function replicates journal entries to the remote system at the Licensed Internal Code layer. Moving the replication to this lower layer provides the following advantages:

- The remote system handles more of the replication overhead
- The overall system performance and journal entry replication performance are improved.
- An option is available to have replication occur synchronously to the operation that is causing the journal entry deposit.
- Journal receiver save operations can be moved to the remote system.

In 7.4.2, “Remote Journal Function” on page 159, you can find a detailed description of the remote journal function.

7.3 Virtual OptiConnect/400

Virtual OptiConnect/400 (see Figure 64 on page 140) uses the same software as OptiConnect/400, but there is no need for any additional hardware units and I/O cards on a system with logical partitions. Virtual OptiConnect emulates the real OptiConnect hardware between logical partitions. In addition, beginning with OS/400 V4R4, you can also run the TCP/IP protocol over both types of OptiConnect.

There are no hardware requirements for virtual OptiConnect. Virtual OptiConnect emulates the real hardware.

If you use virtual OptiConnect, you have the same software requirements as with OptiConnect. You must purchase and install one of the following options:

- 5769SS1 product option 23
- OptiMover PRPQ (5799-FWQ)

Note: You only need a single license of OptiConnect across all partitions. You do not need to purchase OptiConnect with all partitions.

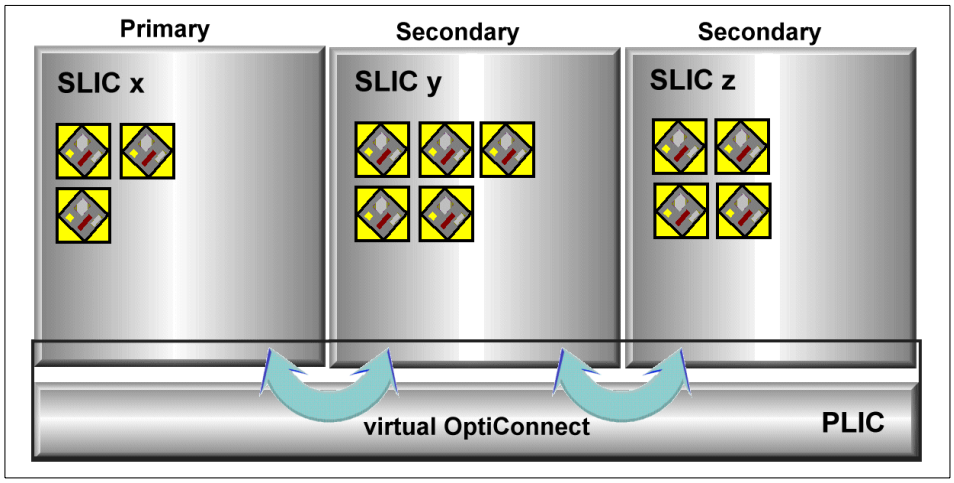


Figure 64. Virtual OptiConnect/400 Concept

If you enabled inter-partition communication using virtual OptiConnect during partition setup, you can communicate with each logical partition.

Beginning with OS/400 V4R4, IBM has significantly reduced the license fee for OptiConnect/400. We recommend that you re-consider the applicability of the OptiConnect/400 solution in your environment again if price was the only issue that prevented you from using this unique technology.

7.3.1 Transferring Data between Logical Partitions

Virtual OptiConnect uses *Direct Memory Access* functions to transfer pages from one partition to another. Figure 65 on page 141 shows an example of a data transfer from partition to partition. The data resides on the DASD in partition 0 and has to be written to the DASD on partition 1. The following process occurs:

1. A read request arrives at the DASD IOP on partition 0. Data is read and transferred through IOP using a Direct Memory Access (DMA) function into memory pages in partition 0.
2. Pages are forwarded, using DMA, into an intermediate buffer on partition 0.
3. To get the data to the memory of partition 1, once again, a DMA is used. Pages are transferred from the intermediate buffer on partition 0 into memory on partition 1. No communication link is used for this *memory copy*. Virtual OptiConnect emulating the real OptiConnect hardware gets

the data from the buffer on partition 0 to the memory of partition 1 by using a virtual OptiConnect IOP.

4. The pages are transferred from the memory of partition 1 through the IOP for the DASD on that partition and written to the disk. This is done again by a DMA function.

Using this kind of mechanism, virtual OptiConnect provides a real high-speed connection between logical partitions. The only likely bottleneck for data transfer is the IOP and DASD arms in the partitions.

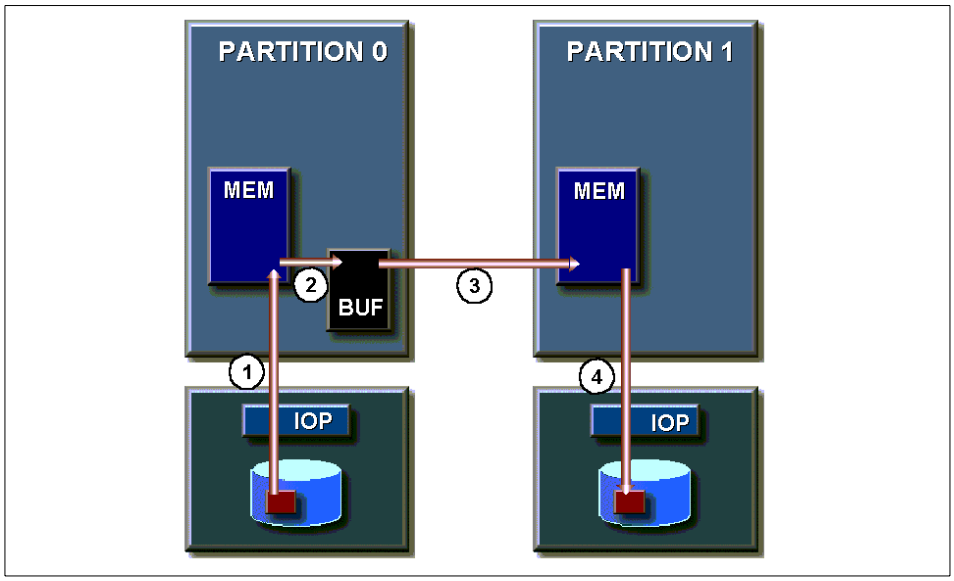


Figure 65. Read Request Using Virtual OptiConnect

7.3.2 Setting Up Virtual OptiConnect over SNA

OptiConnect can be used by any application that was written to use Distributed Data Management (DDM). Many applications that use an AS/400 database can transparently use DDM without changes to the application. Applications that access a database by DDM can use different ways to do so:

- Traditional APPC DDM (SNA)
- OptiConnect/400 (SNA and TCP/IP)
- Virtual OptiConnect (SNA and TCP/IP)

There are two different ways to route data requests through OptiConnect. If you use a special keyword in the DDM file, you will use the Fastpath OptiConnect. Using this method, OptiConnect agent jobs start in the

OptiConnect Connection Manager and run in the QSOC subsystem. Although Fastpath OptiConnect is the faster means of communications, you should be aware that two-phase commit is not supported by this method.

The second method is called Extended Function Path OptiConnect. Using this method, the OptiConnect agent jobs are started by the advanced program-to-program communications (APPC) attach manager and run in the QCMN subsystem. Two-phase commit is supported by Extended Function Path OptiConnect.

7.3.2.1 Setting Up Fastpath OptiConnect

If you install the OptiConnect/400 software, the QYCTSOC device description is created automatically. This device always remains varied off. This device description is necessary for OptiConnect, and therefore, must not be deleted. Whenever an APPC conversation is directed to this device, the OptiConnect device driver redirects the conversation through the OptiConnect bus.

To create a DDM file for Fastpath OptiConnect, perform the following steps:

1. Enter the `CRTDDMF` command and press **F4**.
2. Enter the DDM file name and library. This is the name that the application must know to have access to the data.
3. Enter the remote file name and library. In this library of the remote system, the data file is actually stored. In an environment with logical partitions. this is the name and library for the data file in that logical partition.
4. For the Remote Location parameter, specify the *system name* of the *target system*, where the database file is located. To find out the name, use the Display Network Attributes (`DSPNETA`) command on the remote system. In an environment with logical partitions, this is not the name of the logical partition. It is the name of the system as specified in the logical partition.
5. Press **F10** for additional parameters.
6. On the Additional Parameters display, page down and enter `QYCTSOC` for the APPC device description. Whenever DDM finds a DDM file with this special keyword used in the file description, OptiConnect routes requests using the Fastpath method.
7. In the Mode parameter, you can specify any valid mode of your AS/400 system. If you want to run the OptiConnect Agent jobs to start with USRPRF specified in the QYYCDTSU job description, then you must use the same keyword as on the APPC device description, `QYCTSOC`.

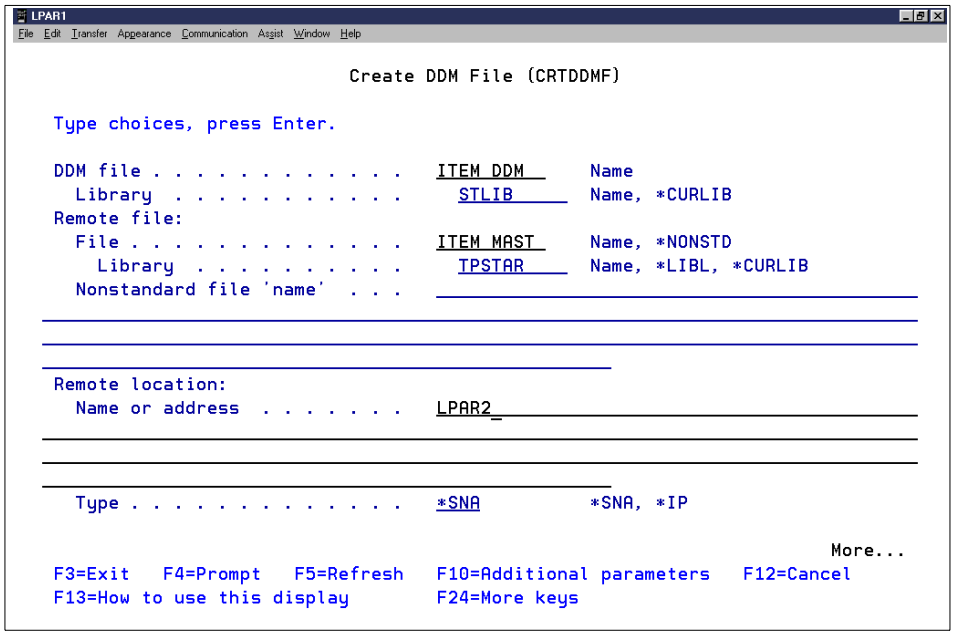


Figure 66. Creating a DDM File to Use OptiConnect (Part 1 of 2)

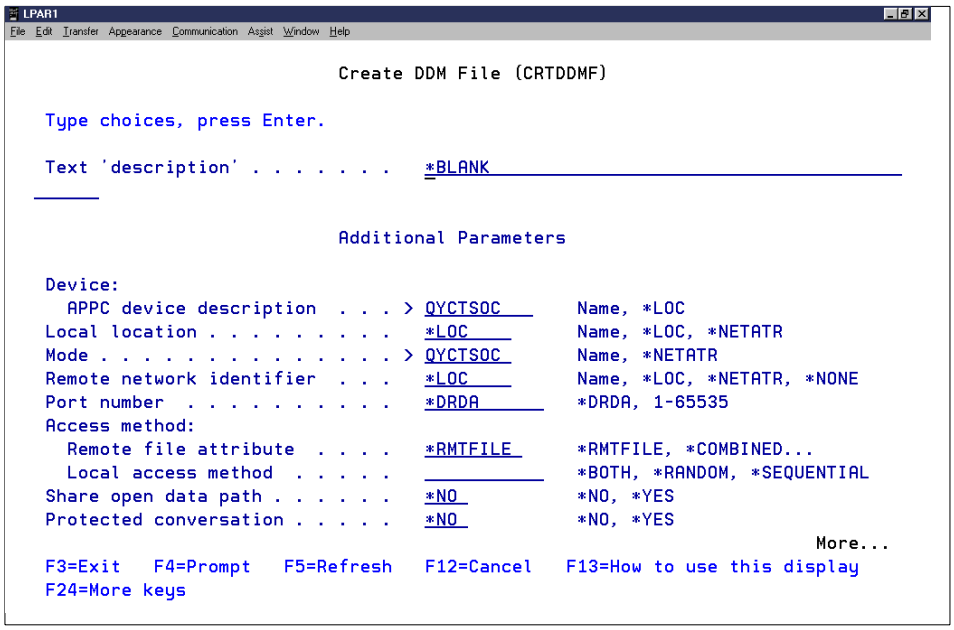


Figure 67. Creating a DDM File to Use OptiConnect (Part 2 of 2)

7.3.2.2 Setting Up Extended Function Path Routing

For such applications as SNA distribution services and the remote journal function, you need to set up Extended Function Path Routing. This enables you to route data requests through OptiConnect without the special device keyword, QYCTSOC. You must create OptiConnect controllers and devices of type *OPC on the logical partitions.

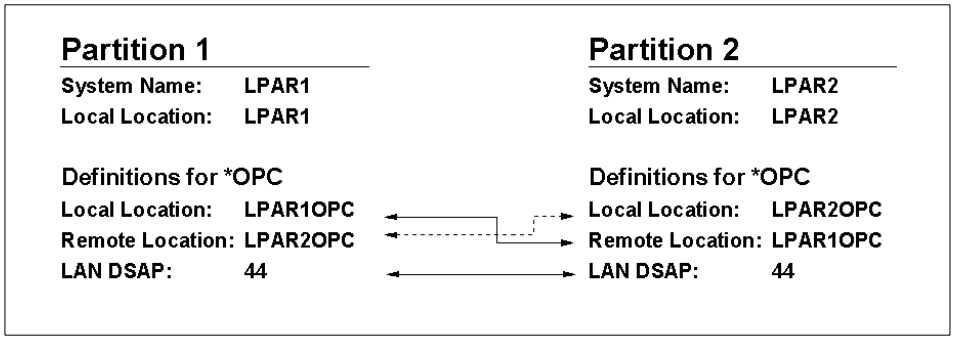


Figure 68. Matching Parameters for Extended Function Path Routing

First, you must create a pair of *OPC controllers, one on each of the logical partitions that uses OptiConnect to communicate to another logical partition. One controller on a logical partition must have *PRI (primary) defined in the Data Link Role. The other controller must be *SEC (secondary). The LAN DSAP parameter value must be valid and identical for both controllers on both partitions in the pair. Valid values are 04, 08, 0C, 10, 14, ... 78, and 7C.

Use the following command to configure the *OPC controller on a logical partition:

```
CRTCTLAPPC CTLD(name) LINKTYPE(*OPC) RMTSYSNAME(sysname) ROLE(*PRI or *SEC)
DSAP(##)
```

Figure 69 and Figure 70 on page 145 show an example of creating two *OPC controllers on partition LPAR1 and on partition LPAR2.

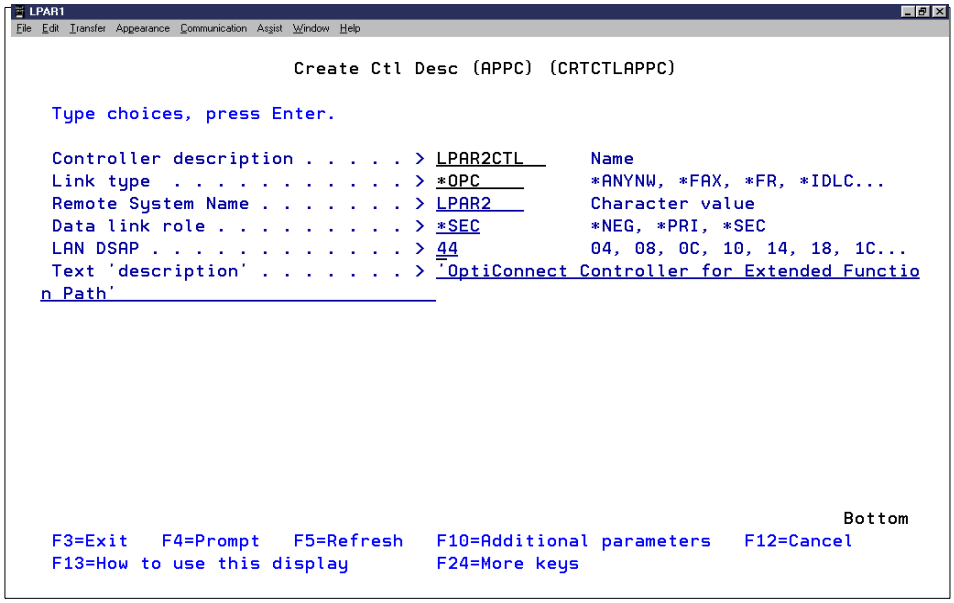


Figure 69. Create Controller Description on LPAR1 to Connect to LPAR2

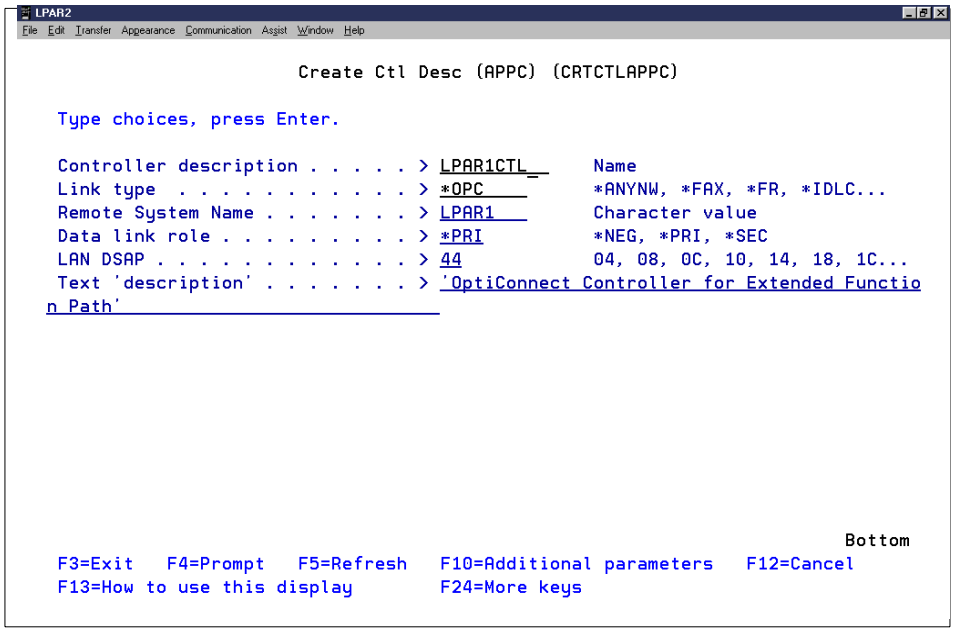


Figure 70. Create Controller Description on LPAR2 to Connect to LPAR1

Now, create device descriptions for each controller on each partition. Specify APPN(*NO), because the *OPC controller only accepts devices that are non APPN-capable.

Use the following command to configure the *OPC controller on a logical partition:

```
CRTDEVAPPC DEVD(name) RMTLOCNAME(LPAR2OPC) ONLINE(*NO)
LCLLOCNAME(LPAR1OPC) APPN(*NO)
```

Figure 71 and Figure 72 on page 147 show an example of creating two device descriptions to attach to the *OPC controller.

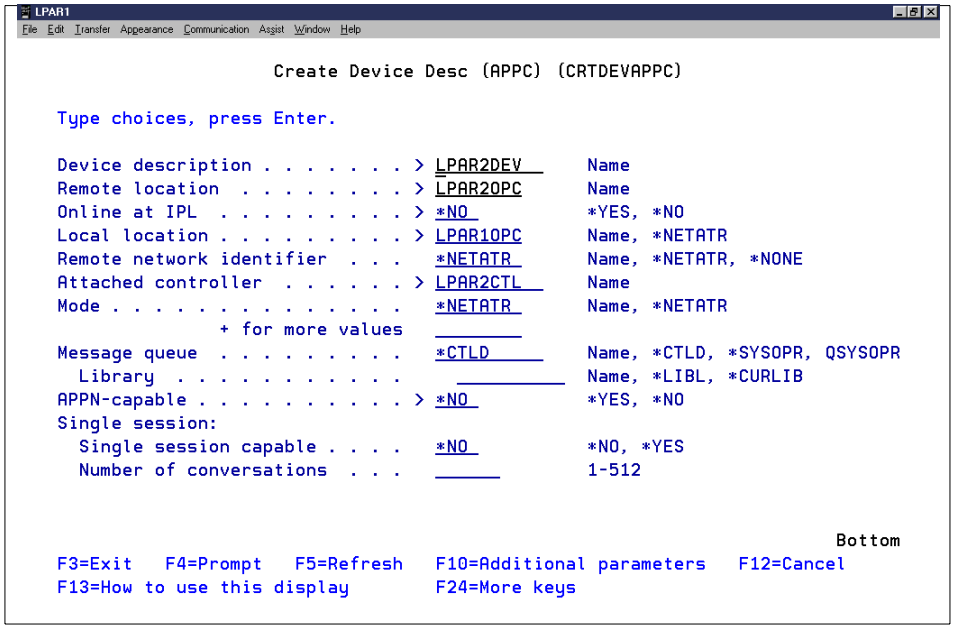


Figure 71. Create Device Description on LPAR1 to Attach LPAR2CTL

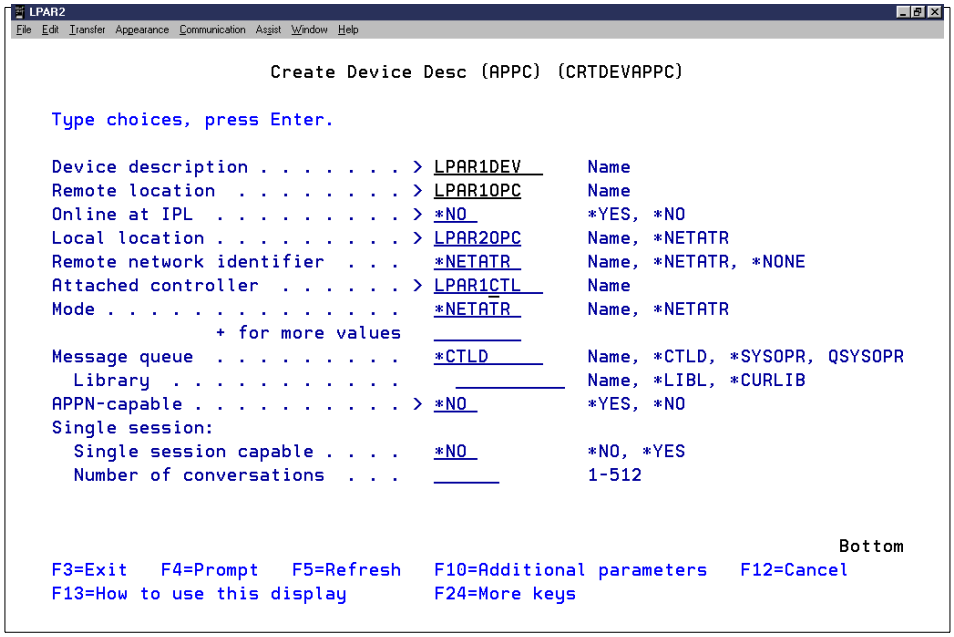


Figure 72. Create Device Description on LPAR2 to Attach LPAR1CTL

These devices do not actually describe a device. They are used to control authority and access to the OptiConnect path.

To enable requests over OptiConnect using the created controllers and devices, you must start the QSOC subsystem first. After starting that subsystem, you can vary on all *OPC controllers and devices. The OptiConnect path is completely established when both controllers and devices change to the ACTIVE state.

These *OPC controllers can be used to provide APPC communication capability over virtual OptiConnect. Some of the applications for that interface are:

- An application program using the inter communication file (ICF) file interface
- The communication programming interface (CPI) communication call interface
- The customer information control system (CICS) file interface

To create a DDM file for Extended Function Path Routing, complete these steps:

1. Enter the CRTDDMF command, and press **F4**.
2. Enter the DDM file name and library. This is the name that the application must know to have access to the data.
3. Enter the remote file name and library. In this library of the remote system, the data file is actually stored. In an environment with logical partitions, this is the name and library for the data file in that logical partition.
4. For the Remote Location parameter, specify the *system name* of the *target system*, where the database file is located. To find out the name, use the Display Network Attributes (DSPNETA) command on the remote system. In an environment with logical partitions, this is not the name of the logical partition. It is the name of the system as specified in the logical partition.
5. Press **F10** for additional parameters.
6. On the Additional Parameters display, page down. Enter QYCTSOC for the APPC device description. Whenever DDM finds a DDM file with this special keyword used in the file description, OptiConnect routes requests using the Fastpath method.

In the Mode parameter, you can specify any valid mode of your AS/400 system. If you want to run the OptiConnect Agent jobs starting with USRPRF specified in the QYYCDTSU job description, then you must use the same keyword as on the APPC device description, QYCTSOC.

7.3.3 Starting Virtual OptiConnect

To start virtual OptiConnect, you have to start subsystem QSOC. The best way to do this, is by adding a command to your startup program that will be run at IPL time.

To start the QSOC subsystem, you have to enter the Start Subsystem (STRSBS) command on each partition:

```
STRSBS QSOC/QSOC
```

When you start the QSOC subsystem, the OptiConnect connection manager, SOCMGR, starts as an autostart job.

7.3.4 Ending Virtual OptiConnect

You can end OptiConnect by ending the QSOC subsystem. Before you end OptiConnect on a logical partiton, make sure that there are no application programs still using this connection.

If you use the remote journal function on this partition, end this function before ending the QSOC subsystem. remote journal function jobs are not displayed with the Work with Active Jobs (WRKACTJOB) command.

To end virtual OptiConnect, enter the following command:

```
ENDSBS QSOC *IMMED
```

If you ended the QSOC subsystem on one logical partition, you still can use virtual OptiConnect on your system if you have more partitions that have enabled Interpartition OptiConnect.

If you configured Extended Function Path Routing by creating controllers and devices, vary off the *OPC controllers and the corresponding controllers on the other logical partition. These controllers are in an unusable state when you end the QSOC subsystem. You have to vary off and then vary on those controllers to activate them.

7.3.5 Determining Virtual OptiConnect Activity

Use the following commands to verify the OptiConnect activity and to obtain information about its resources and components.

- Work with Active Jobs (WRKACTJOB)
- Work with OptiConnect Activity (WRKOPCACT)
- Display OptiConnect Link Status (DSPOPCLNK)

7.3.5.1 Work with Active Jobs (WRKACTJOB)

You can monitor OptiConnect activity by using the Work with Active Jobs (WRKACTJOB) command. To see the active jobs running in the QSOC subsystem, enter the following command:

```
WRKACTJOB SBS(QSOC)
```

You will see a SOCMGR job if the QSOC subsystem is running. On the target partition, you will also see one or more agent jobs (SOCAnnnnnn jobs) running in susbsytem QSOC.

Agent jobs are active or inactive (prestarted). To determine if an agent job is active, enter a `5` in front of the agent job in the WRKACTJOB display. Then choose option **11** (Display call stack) or option **14** (Display open files). Inactive agent jobs have no files open.

Active agents jobs are present until one of the following actions occur:

- The source partition job ends or the user logs off.
- The source partition job ends or the user uses the Reclaim Resources (RCLRSC) command.
- The source partition job ends or the user uses the Reclaim DDM Conversations (RCLDDMCNV) command.

Figure 73 shows a sample of the Work with Active Jobs display.

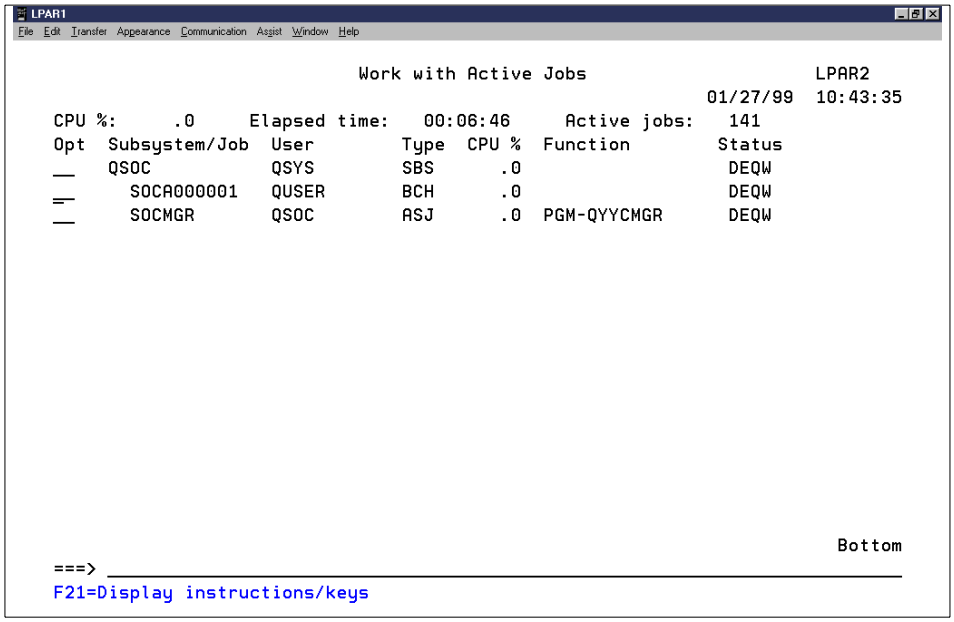


Figure 73. Work with Active Jobs — QSOC Subsystem

7.3.5.2 Work with OptiConnect Activity (WRKOPCACT)

This is the most useful command in an OptiConnect environment. Using the Work with OptiConnect Activity (WRKOPCACT) command, you get information about the bus activity. You can also change the state of OptiConnect resources. Plus, you can see information about database transactions using the virtual OptiConnect bus and the connection status of all involved client and server partitions.

The Work with OptiConnect Activity provides three different views:

- The first view you see when you enter the WRKOPCACT command. It shows the OptiConnect connectivity from the perspective of an application partition. Here is the only place where you can see the number of transactions between logical partitions. You see the connection status of virtual OptiConnect to any other logical partition. The collection period is shown on the top of the display.
- The next view shows this partition as a client. You can access this view by pressing **F11** (Client Statistics View).
- The next view shows this partition as a server. You can access this view by pressing **F11** (Server Statistics View).

Figure 74 shows the Work with OptiConnect Activity when you enter the following command:

```
WRKOPCACT
```

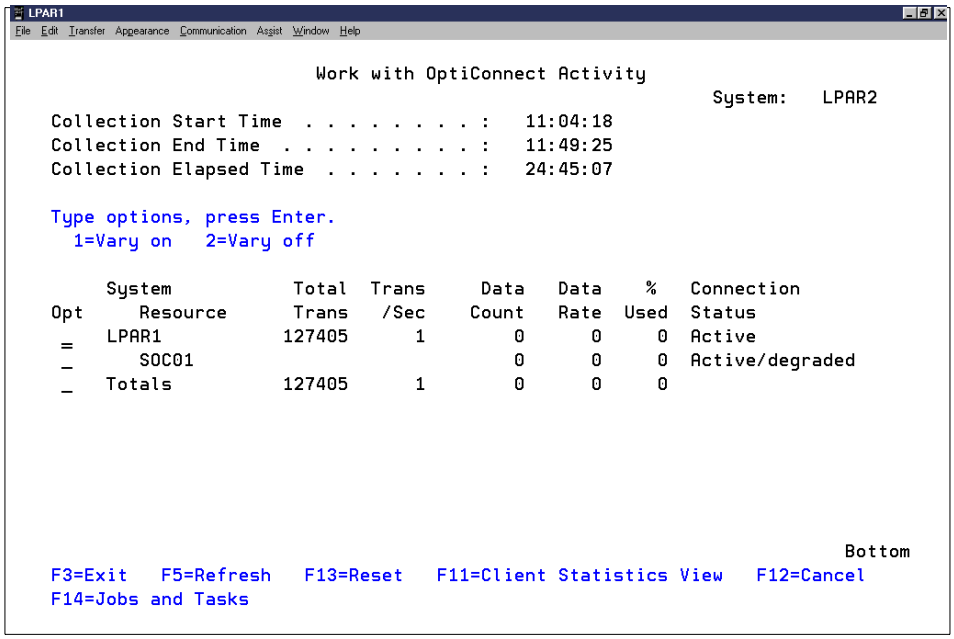


Figure 74. Work with OptiConnect Activity

The *Active/degraded* connection status indicates that this system has an operational connection to the system named in the System column through the adapter resource. This connection is in use by user or agent jobs.

The degraded status means redundancy is reduced because this resource has one or more unusable local or remote links. To determine the reason for the resource degraded status, check the status of both the local and remote links. A link status other than Ready or Active indicates that the link is unusable and will result in the degraded status. A blank connection status indicates that the link does not exist and will not show the degraded status.

7.3.5.3 Display OptiConnect Link Status (DSPOPCLNK)

The Display OptiConnect Link Status (DSPOPCLNK) command shows the connection status between systems in the fiber-optic network. In an environment with logical partitions, you can use this command to check the status of the connection. Most of the other views of this command do not apply to virtual OptiConnect. There are no real OptiConnect hardware resources in a virtual OptiConnect Environment.

The Top Link, for example, shows the fiber optic connection that runs from the top port of the System Unit Optical Link Processor to the top port of the OptiConnect Adapter. The Bottom Link shows the fiber optic connection that runs from the bottom port of the System Unit Optical Link Processor to the bottom port of the OptiConnect Adapter. The Redundant Link shows the fiber optic connection that runs from the top port of one OptiConnect Adapter to the bottom port of the other OptiConnect Adapter. In an environment with virtual OptiConnect, the top link and the bottom link are shown as Unknown and the Redundant link is left blank.

7.3.6 Setting Up OptiConnect/400 over TCP/IP

Prior to OS/400 Version 4 Release 4, you could only use the SNA protocol to communicate over OptiConnect. Beginning with V4R4, you can also have applications that use Transport Control Protocol/Internet Protocol (TCP/IP) to communicate over OptiConnect. The AS/400 provides a standard IP interface, which allows existing applications and services to work unchanged by defining a TCP/IP interface that uses OptiConnect.

The IP over OptiConnect implementation allows applications that use a TCP/IP Socket interface to communicate over the OptiConnect shared bus when running in an AS/400 cluster. This allows the growing set of applications that are distributed across multiple AS/400 systems (three-tiered applications) to take advantage of the high bandwidth and low latency of the OptiConnect shared bus. Prior to this implementation, the OptiConnect shared bus was available using a proprietary API (OptiMover) or using an APPC controller. The TCP/IP socket interface to OptiConnect provides the highest bandwidth, with the lowest latency communication method between AS/400 systems and between logical partitions. The IP over OptiConnect

implementation provides a *line driver* interface between the bottom of the TCP/IP stack and the OptiConnect device driver. It provides communication services over a shared bus in an AS/400 cluster.

You can configure up to four IP interfaces to be used by OptiConnect. Each of the interfaces must be on a separate subnet. To configure this interface, you can use the standard AS/400 methods, which are Configure TCP/IP (CFGTCP) and the Operations Navigator graphical user interface (GUI).

Once you configure the IP interface, you have to start it. It allows you to communicate to other partitions that have enabled Inter-partition OptiConnect. To do so, those partitions must have an interface with the same subnet. All protocols which use Internet Protocol (IP), including Transmission Control Protocol (TCP), User Datagram Protocol (UDP), and so forth.

7.3.6.1 Defining the OptiConnect Interface

To add a new interface to the Transmission Control Protocol/Internet Protocol (TCP/IP), you can use the Add TCP/IP Interface (ADDTCPIFC) command. If you add a new interface, which will be used by virtual OptiConnect, you do not have any hardware resources that you are really using. Therefore, the interfaces created by the ADDTCPIFC command for virtual OptiConnect are logical interfaces.

7.3.6.2 Configuring the OptiConnect Interface

You can configure the TCP/IP interfaces for OptiConnect in two different ways. In the first configuration, the OptiConnect bus has to be seen similar to a LAN.

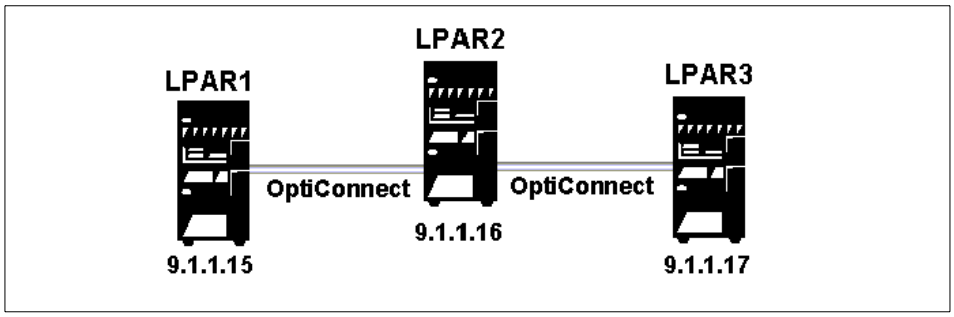


Figure 75. Configuration of the OptiConnect Interface — LAN Version

In Figure 75, you see an example for an environment with logical partitions. We have three logical partitions and all are Inter-partition OptiConnect enabled. To configure the interface for use by OptiConnect, use the Add

TCP/IP Interface (ADDTCPIFC) command. Each *OPC interface has to be assigned a unique IP address within the subnet. Even if it seems that there are two interfaces for LPAR2, it is only one interface with the IP address *9.1.1.16*. This interface is attached to the optical bus (OptiConnect). You will see a new parameter special value here, *OPC. This special value is used if you are adding an OptiConnect interface over TCP/IP.

The required parameters are:

INTNETADR (Internet address)
Specifies an Internet address that the local system responds to on this interface. An interface is associated with a line description. The Internet address is specified in the form nnn.nnn.nnn.nnn, where nnn is a decimal number ranging from 0 through 255. An Internet address is not valid if it has a value of all binary ones or all binary zeros for the network identifier (ID) portion or the host ID portion of the address. If the Internet address is entered from a command line, the address must be enclosed in apostrophes.

LIND (Line description)
Specifies the name of the line description associated with the new interface. The line description must be defined before the TCP/IP interface can be added. In an OptiConnect environment, you must use the special value *OPC here.

SUBNETMASK (Subnet mask)
Specifies the subnet mask, which is a bit mask that defines the part of the network where this interface attaches. The mask is a 32-bit combination that is logically ANDed with the Internet address to determine a particular subnetwork. The bits of the mask set to the value one (1) determine the network and subnetwork portions of the address. The bits set to the value zero (0) determine the host portion of the address.

LCLIFC (Associated local interface) — *Not a required parameter*
Use this parameter to associate the interface that you are currently defining with an existing local TCP/IP interface. The associated local interface (LCLIFC) is used to allow *transparent subnetting* (also known as *Proxy Arp*) between the associated interfaces, to define Frame Relay unnumbered networks or for load balancing.

To configure the TCP/IP interfaces for this example, enter the following command statement on partition LPAR1:

```
ADDTCPIFC INTNETADR('9.1.1.15') LIND(*OPC) SUBNETMASK('255.555.255.0')
```

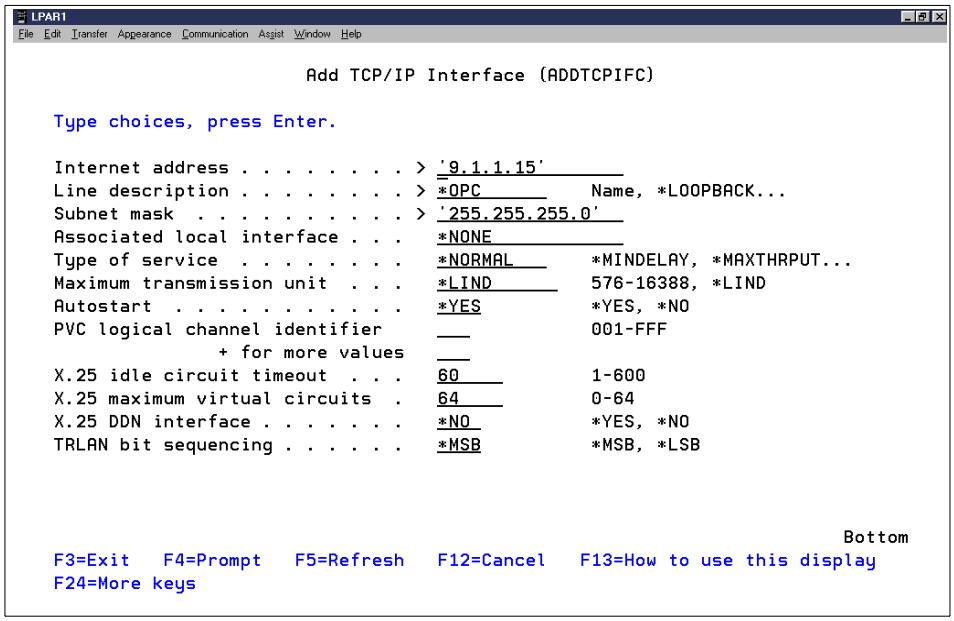


Figure 76. ADDTCPIFC — Add TCP/IP Interface for OptiConnect

To configure the interfaces on LPAR2 and LPAR3, enter the following command statement for partition LPAR2:

```
ADDTCPIFC INTNETADR('9.1.1.16') LIND(*OPC) SUBNETMASK('255.555.255.0')
```

For partition LPAR3, enter this command statement:

```
ADDTCPIFC INTNETADR('9.1.1.17') LIND(*OPC) SUBNETMASK('255.555.255.0')
```

Another approach for configuring TCP/IP is to use the associated local interface parameter. If you use this method, there is no reason to define new subnets for the OptiConnect bus. No external route tables need to be updated to provide connectivity between the OptiConnect interfaces and the rest of the TCP/IP network. You use the OptiConnect interfaces as part of existing local subnets to which the AS/400 system is attached.

Each OptiConnect interface defines an endpoint of a point-to-point OptiConnect connection between two AS/400 systems. Configuring the interface configured this way is actually not an interface in the system or partition to which you are currently signed on. Instead, it is an existing interface in the endpoint of that connection. The existing local interface for the TCP/IP network will be specified as the associated local interface to the OptiConnect interface.

In Figure 77, you see an environment with three logical partitions. All of these partitions have their own interface to the Token-Ring network.

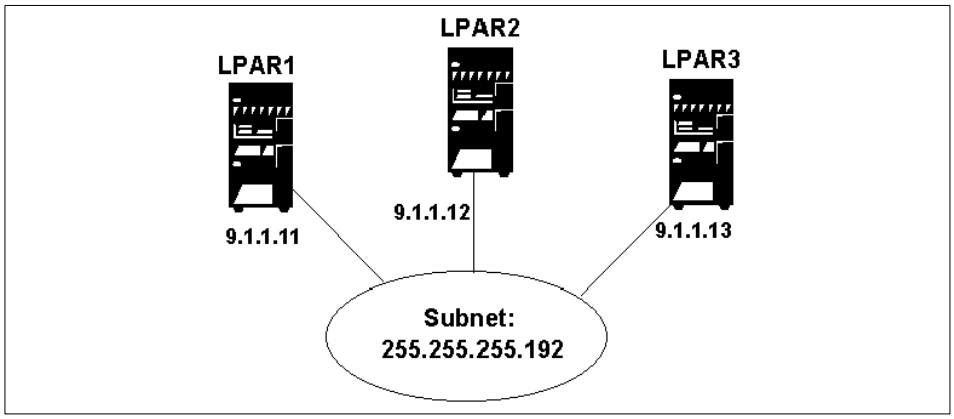


Figure 77. Configuring OptiConnect Interface (Part 1 of 2)

Logical partition LPAR1 has the Internet address `9.1.1.11`, LPAR2 has the Internet address `9.1.1.12`, and LPAR3 has the Internet address `9.1.1.13`. The subnet mask for these three partitions is `255.255.255.192`.

Now we want to add the OptiConnect interfaces for this example. You can see the new configuration in Figure 78.

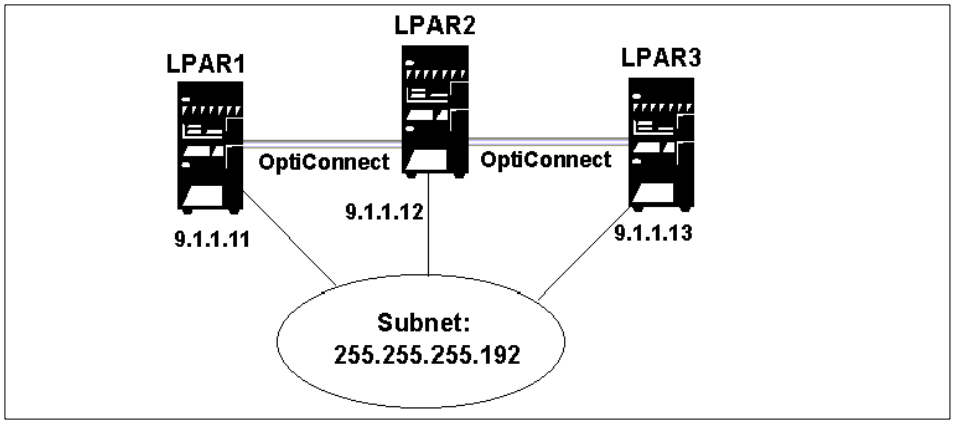


Figure 78. Configuring OptiConnect Interface (Part 2 of 2)

On partition LPAR1, perform the following steps:

1. Define the endpoint of the OptiConnect connection for logical partition LPAR2 using the associated local interface: `9.1.1.11`

   The Internet address for the endpoint in logical partion LPAR2 is: `9.1.1.12`

2. Define the endpoint of the OptiConnect connection for logical partition LPAR3 using the associated local interface: `9.1.1.11`

   The Internet address for the endpoint in logical partition LPAR3 is: `9.1.1.13`

Use the following commands for the configuration in LPAR1:

1. For the endpoint in partition LPAR2, enter the following command statement:

```
ADDTCPIFC INTNETADR('9.1.1.12') LIND(*OPC)
SUBNETMASK('255.555.255.255') LCLIFC('9.1.1.11')
```

2. For the endpoint in partition LPAR3, enter the following command statement:

```
ADDTCPIFC INTNETADR('9.1.1.13') LIND(*OPC)
SUBNETMASK('255.555.255.255') LCLIFC('9.1.1.11')
```

No local interfaces for the OptiConnect bus have to be configured. The subnet mask in this configuration method *must* always be `255.255.255.255` for the interfaces describing the endpoint of an OptiConnect connection.

You need to perform these configurations on partition LPAR2 and on partition LPAR3. From the view of LPAR2, you have to configure the endpoints in partitions LPAR1 and LPAR3. From the view of partition LPAR3, you have to configure the endpoints in partitions LPAR1 and LPAR2.

Use the following commands for the configuration in LPAR2:

1. For the endpoint in partition LPAR1, enter:

```
ADDTCPIFC INTNETADR('9.1.1.11') LIND(*OPC)
SUBNETMASK('255.555.255.255') LCLIFC('9.1.1.12')
```

2. For the endpoint in partition LPAR3, enter:

```
ADDTCPIFC INTNETADR('9.1.1.13') LIND(*OPC)
SUBNETMASK('255.555.255.255') LCLIFC('9.1.1.12')
```

Use the following commands for the configuration in LPAR3:

1. For the endpoint in partition LPAR1, enter:

```
ADDTCPIFC INTNETADR('9.1.1.11') LIND(*OPC)
SUBNETMASK('255.555.255.255') LCLIFC('9.1.1.13')
```

2. For the endpoint in partition LPAR2, enter:

```
ADDTCPIFC INTNETADR('9.1.1.12') LIND(*OPC)
SUBNETMASK('255.555.255.255') LCLIFC('9.1.1.13')
```

Another advantage of this method is that, if one of the OptiConnect paths becomes inactive, packets are automatically routed over the backup interface.

If subsystem QSOC has been started and the TCP/IP interfaces are active, all the traffic between the logical partitions are routed over the OptiConnect path.

7.4 Application Examples Using Virtual OptiConnect/400

If you configured your AS/400 system for using logical partitions, you have the capability to use virtual OptiConnect without needing new hardware. You need OS/400 product option 23 (OptiConnect), and you need to enable the Inter-partition OptiConnect during configuration of the logical partitions. These are the hardware and software requirements for the application examples described in the following sections.

7.4.1 SNA Distribution Services (SNADS)

If you want to use the OptiConnect communications path between AS/400 logical partitions for SNA Distribution Services, you have to configure *APPC controllers. You can find information on how to configure *APPC controllers and devices in 7.3.2.2, “Setting Up Extended Function Path Routing” on page 144.

You need to create a directory entry, a routing table entry, and a distribution queue. The definition of the distribution queue has to use the names of the local location and the remote location as defined in the *APPC device description.

7.4.1.1 Adding a Directory Entry

You need to add an **ANY* entry for the logical partition in your directory. To add this entry, enter this command statement:

```
ADDDIRE USRID(xxx/*ANY) USRD(yyy) SYSNAME(zzz)
```

In this statement, note the following explanation:

- xxx is the name of the remote system.
- yyy is the user description.
- zzz is the name of the remote system.

7.4.1.2 Adding a Distribution Queue

You need to add a distribution queue for the logical partition. To add the queue, enter the following command statement:

```
ADDDSTQ DSTQ(xxx) RMTLOCNAME(yyy)
```

Note the following explanation:

- xxx is the name of the queue.
- yyy as specified in the *APPC device.

Then, enter the following statement:

```
DSTQTYPE(*SNADS)
MODE(*NETATR)
```

Or, specify a mode.

```
RMTNETID(*NETATR)
LCLLOCNAME(zzz)
```

In this statement, zzz as specified in the *APPC device.

7.4.1.3 Creating a Routing Table

You need to add a routing table for the logical partition. Enter this statement:

```
ADDDSTRTE SYSNAME(xxx)
```

Here, xxx is the name of the remote system. Note these options:

```
FAST(yyy)
STATUS(zzz)
DATAHIGH(aaa)
DATALOW(bbb)
```

7.4.2 Remote Journal Function

The remote journal function was introduced in OS/400 with V4R2. This function offers a fast and reliable method to transfer journal receiver data to a remote AS/400 system. With the availability of logical partitions on the AS/400 system, the remote journal function is ideal for data replication in a logical partition environment.

You can use the remote journal function with application programs to maintain a replica of the local database. A replica database is a copy of the original

database that resides on another system or on another partition. Before the introduction of the remote journal function, you had to use the Receive Journal Entry (RCVJRNE) command to receive journal entries. A RCVJRNE exit program then receives these entries from the journal receiver and sends them to the remote system or logical partition. This could be done by any available communications method.

For more information on this function, refer to the redbook *AS/400 Remote Journal Function for High Availability and Data Replication*, SG24-5189.

7.4.3 Data Replication Using Journals

There are a lot of high availability solutions in the marketplace using journals for data replication. Solutions that are available from IBM Business Partners, use local journaling and the Receive Journal Entry (RCVJRNE) command. In this environment, database changes are generated by user's applications running on a source production system. These database changes create journal entries in a local journal receiver. This is a common functionality used by applications on AS/400 systems.

These journal entries in a local journal receiver must be transferred to the remote system. The high availability solutions provided by IBM business partners use the Receive Journal Entry (RCVJRNE) command and an exit program to transfer the journal entries to the target system. The communications transport can be done using one of the available communication lines to the target system. The fastest communication method is virtual OptiConnect. The data is transmitted asynchronously to the target system. On the target system these journal entries are retrieved and usually stored in a user space. Another job, or many jobs, replay the changes into a copy of the source database. Once the journal entries are applied to the database, you have an exact copy of the production database on the target system.

The disadvantage of this solution is:

- The asynchronous transfer of journal receiver entries means data latency. If your source systems fails while there are some journal entries still waiting for transmission, a few final database transactions are trapped on the source system. The database replica on the target system lags behind the production database on the source system.
- The data passes through many layers of the system software. It crosses the Machine Interface (MI) boundary several times. This means extra CPU cycles and therefore increased CPU utilization.

In Figure 79, you can see a high-level description of this process.

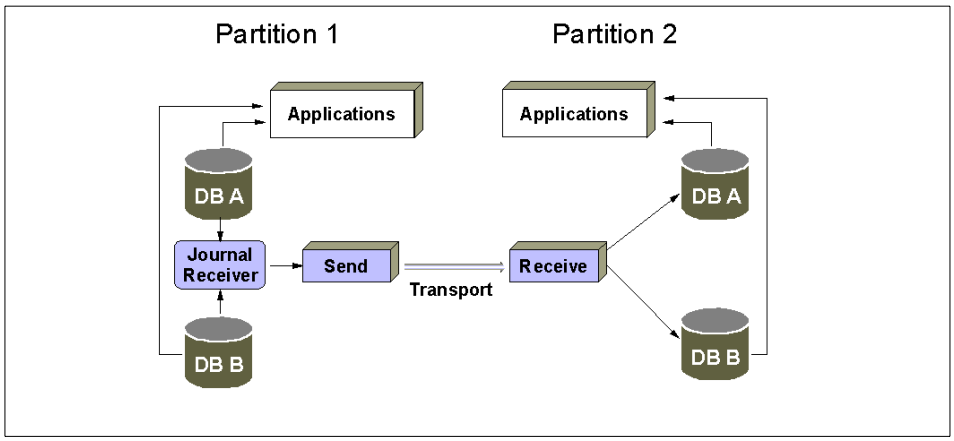


Figure 79. Hot-Backup Environment without Remote Journal Function

7.4.4 Data Replication Using the Remote Journal Function

To get rid of the disadvantages, data latency and increased CPU utilization, the remote journal function provides a much more efficient method of journal entry transport. Efficient low-level system code is used to capture and transmit journal entries directly from the source system to associated journals and journal receivers on a target system. There is no need to buffer the resulting journal entries on the production system. Much of the processing is done below the MI. The CPU utilization decreases using this method. If the remote journal function was activated in synchronous mode, the journal entries are replicated to the target system's main storage before the production system's database gets updated. The data latency is driven to zero.

There are still more tasks in a high availability solution on the AS/400 system than the replication of the database. In fact, the available high availabilty solutions can take advantage of this more efficient transport mechanism. In customer environments for a hot-backup scenario, there is a need for providing the required management facilities. The available solutions for hot-backup support replication for more system objects than the database.

In Figure 80 on page 162, you can see a high-level figure for a hot-backup solution using the remote journal function.

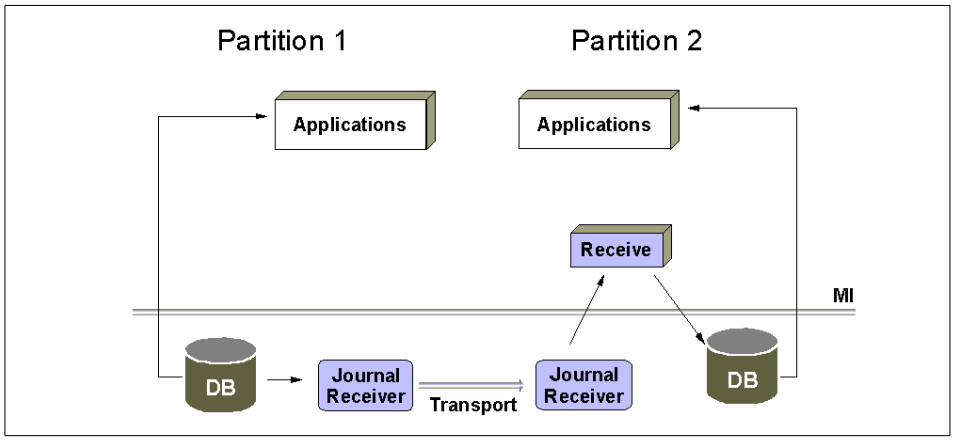


Figure 80. Hot-Backup with Remote Journal Function

7.4.5 Setting Up the Remote Journal Function

The remote journal function is built upon the local journal function. Before the remote journal function can be performed, the local journal function on the source system must first be set up.

Complete the following steps to setup the remote journal function:

1. Set up a relational database (RDB) directory entry for the remote system. The RDB directory contains database names and values that are translated into communications network parameters. In an RDB directory entry, specify the communications protocol and the corresdonding communications path. In Figure 81 on page 163, you can see an example of how to add an RDB entry using virtual OptiConnect.

   You must add one entry for each remote RDB. The RDB must be unique within the RDB directory and within the distributed network.

```
                    Add RDB Directory Entry (ADDRDBDIRE)

Type choices, press Enter.

Relational database  . . . . . > LPAR2
Remote location:
  Name or address  . . . . . > LPAR2OPC



  Type . . . . . .   *SNA          *SNA, *IP
Text . . . . . > 'Remote RDB on Partition LPAR2'

Device:
  APPC device description  . . . > LPAR2OPC      Name, *LOC
Local location . . . . . > PRIMEOPC      Name, *LOC, *NETATR
Remote network identifier  . . . > *NONE         Name, *LOC, *NETATR, *NONE
Mode . . . . . .   *NETATR       Name, *NETATR
Transaction program  . . . . . .   *DRDA         Character value, *DRDA

                                                                        Bottom
F3=Exit   F4=Prompt   F5=Refresh   F12=Cancel   F13=How to use this display
F24=More keys
```

Figure 81. Add RDB Directory Entry — SNA

2. If you use TCP/IP over OptiConnect for the transfer of the remote journalentries, you have to change the DDM server job attributes for TCP/IP. The DDM server should start on IPL.

   You need to specify whether client systems are required to have a password in addition to a user ID on incoming connection requests to this partition as a server. This parameter takes effect on the next DRDA or DDM connect request over TCP/IP.

```
                  Change DDM TCP/IP Attributes (CHGDDMTCPA)

Type choices, press Enter.

Autostart server . . . . . .   *YES          *NO, *YES, *SAME
Password required  . . . . . .   *YES          *NO, *YES, *ENCRYPTED, *SAME
```

Figure 82. Change DDM TCP/IP Attributes

3. Create the local journal and the local journal receiver on the production system by using the Create Journal Receiver (CRTJRNRCV) command and the Create Journal (CRTJRN) command.

4. Start local journaling using the Start Journal Physical File (STRJRNPF) command.
5. Add the remote journal using the Add Remote Journal (ADDRMTJRN) command. This command associates the remote journal on the target system with the specified journal on the source system.
6. Before journal entries can be replicated to a remote journal, the remote journal must be activated. To activate a remote journal, use the Change Remote Journal (CHGRMTJRN) command. This command *must* be run on the source system, on which the source journal is located. Each remote journal must be activated individually by using this command.

Using the Change Remote Journal (CHGRMTJRN) command, you decide which mode you use for the replication of the journal entries.

7.4.6 Benefits of the Remote Journal Function

The remote journal function is well suited on a logical partition environment. You can offload workload from the production partition to the backup partition.

The remote journal function replicates journal entries to the remote partition at the Licensed Internal Code layer. Moving the replication to this lower layer provides these advantages:

- The remote partition handles more of the replication overhead.
- The overall system performance and journal entry replication performance is improved.
- An option is available to have replication occur synchronously to the operation that causes the journal deposit. Using this mode, journal entries are replicated to the main memory on the remote partition first. After an arrival confirmation is returned to the source partition, the journal entry is deposited to the local receiver. Next, the actual database update is made.
- Journal receiver save operations can be moved to the remote system.

7.4.7 Using Your Database 24 Hours a Day, 7 Days a Week

If you need to provide 24-hour by 7-day operation of your AS/400 system, you have to consider a solution that gives you the ability to run your business as well as take care of the housekeeping of your vital data. Figure 84 on page 166 shows an example of how the remote journal function can help in such an envirnonment.

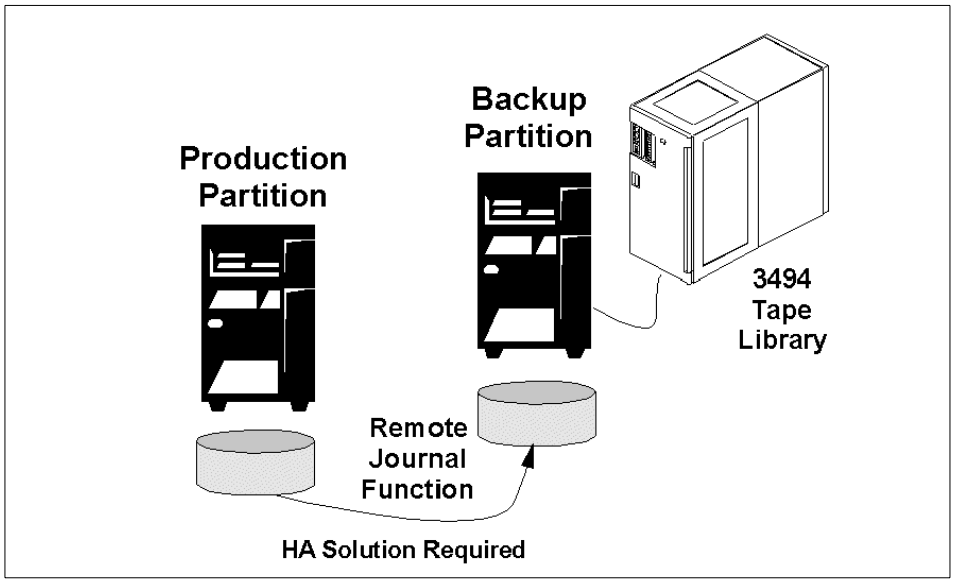


Figure 83. Using Remote Journal Function for 24-by-7 Database Availability

The journal entries are transferred to the backup partition using the remote journal function. The changes are then applied to the backup copy of the application so it reflects what is on the production partition. It is then possible to save from the backup partitions copy of the application, which would ensure that the production partition is available to the end users.

7.4.8 DB2 Multisystem

DB2 Multisystem is a parallel processing technique that provides an almost unlimited scalability option for databases. Using DB2 Multisystem, you have the capability to attach multiple AS/400 systems (up to 32 systems) together in a "shared nothing" cluster. *Shared nothing* means that each system in the coupled network owns and manages its own main memory and disk storage.

Once the AS/400 systems are connected, database files can be spread across the storage units on each connected system. To users, the file looks like a local file on their system. From the user's perspective, the database appears as a single database. The user can run queries in parallel across all the systems in the network and have realtime access to the data in the files.

To install DB2 Multisystem, use option **27** in the list of installable options for the OS/400 operating system.

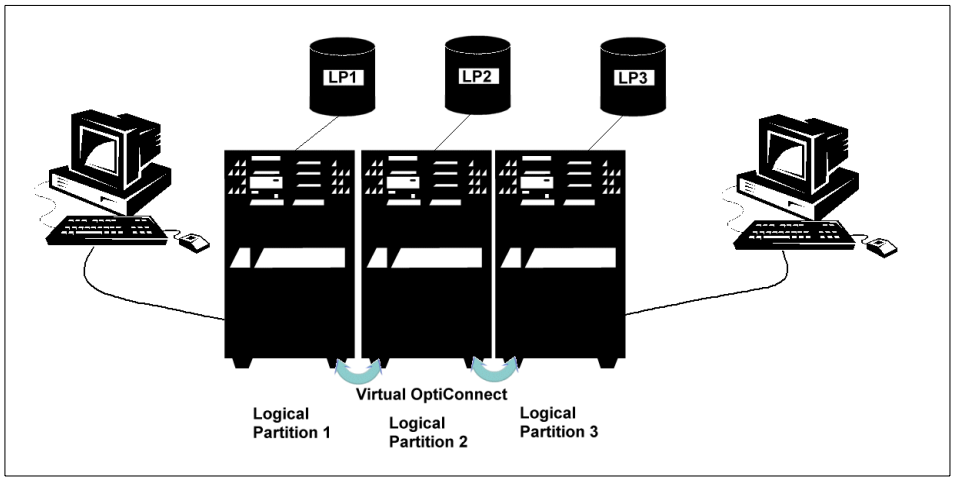


Figure 84. Distribution of Database Files across Logical Partitions

In Figure 84, the database is spread across three logical partitions. Running a query against a database file allows each partition that contains a part of the database to run its portion of the query at the same time as the other partitions. Once each partition has run its part of the query over its local data, the results are combined and returned to the user. Each partition processes a smaller amount of data, which results in the overall query completing in a much shorter amount of time.

Appendix A. Logical Partitioning Terminology

This appendix reviews most of the common hardware terms used within the logical partitioning context.

A.1 Processor

All machines have at least one processor that drives the primary partition. Through logical partitioning, multi-processor (N-way) machines may be configured with secondary partitions. Each partition contains one or more processors.

A.2 Main Storage

Main storage is comparable to RAM on a personal computer. Sufficient main storage must be allocated to primary and secondary partitions to ensure optimal performance.

A.3 Bus

A bus is a piece of hardware used by the AS/400 system to bind its individual physical parts into a single machine. It is like a highway used by the processor to gain access to all its parts. Within the logical partitioning function, buses are some of the components allocated to logical partitions to create independent systems within the physical machine.

Resources can be allocated at a bus level. This means that all resources attached to the bus are allocated to a single bus at a time.

A.4 Input/Output Processor (IOP)

The input/output processor (IOP) is a piece of hardware that allows devices, such as disks and tape drives, or communications lines, to attach to a bus, using an input/output adapter (IOA). They connect to the system like this: the device is connected to the IOA, which is connected to the IOP, which is connected to the bus.

Resources can be allocated at an IOP level. In this case, the bus to which the IOP is attached is shared, and all the resources attached to the IOP are allocated to a single partition at a time. It is then possible to switch this IOP from partition to partition as the need arises.

A.5 Multifunction IOP (MFIOP)

This is a type of input/output processor (IOP) that can connect to a variety of different input/output adapters (IOA). For example, an MFIOP can support disk units, a console, and a tape drive.

There is a second type of multifunction IOP called a *combined multifunction IOP (CFIOP)*. Conceptually, they are the same. However, there are some special rules on how they attach to a bus and are allocated to logical partitions.

A.6 Load Source

The *load source* is a special disk that contains the base software for the machine, called *System Licensed Internal Code (SLIC)*. It also contains information describing how the logical partitions are configured.

Logical partitions represent independent systems. As such, each partition must have a Load Source disk containing its own System Licensed Internal Code and logical partitioning configuration data. One of the actions required in the creation of logical partitions is to define the disk unit which will become the Load Source.

A.7 System Licensed Internal Code (SLIC)

The System Licensed Internal Code or SLIC is the microcode that drives the hardware. On a partitioned machine, each partition has its own SLIC loaded on to the Load Source. This allows partitions to run on different levels of system software if necessary.

A.8 Partition Licensed Internal Code (PLIC)

The Partition Licensed Internal Code (PLIC) is a thin layer of microcode, which is common to all logical partitions in a partitioned machine. It contains the code required to implement logical partitioning across the physical machine.

A.9 Hypervisor

The Hypervisor is made of a small portion of the primary partition SLIC and the PLIC, which runs across all partitions. It is the enabling agent for logical partitioning.

A.10 Removable Media Device

Every logical partition must have either a tape drive or an optical CD-ROM device. This is used to load software (for example, the System Licensed Internal Code) and data into the partition. Such devices are known as *removable media devices*.

A.11 Console

Each logical partition must have a console attached to it through an IOP. Through this console, each logical partition looks and functions like an independent system. The console allows you to enter commands and view the operation of a logical partition.

A primary partition console allows you to see and manage many aspects of all secondary partitions. A secondary partition console allows you access only to its attached partition.

A.12 Expansion Unit

Expansion units are racks attached to a system to accommodate features and devices external to the system unit. Some expansion units can support disk only. Others can support a variety of IOPs, including disk and tape controllers.

The IOPs on an expansion unit access the system unit using a bus that connects the expansion unit to the system unit.

Expansion units may be needed in the creation of logical partitions to provide support for all of the resources required.

A.13 Ownership and Sharing

Logical partitioning allows the switching of some resources between partitions. This can be done at the bus level or the IOP level.

A bus can be owned in two ways: *dedicated* or *shared*. In a dedicated mode, the bus belongs exclusively to one partition. In a shared mode, the IOPs attached to the bus can be switched from partition to partition.

The permutations used for these terms are:

- Own bus dedicated (indicating a non-shareable resource)
- Own bus shared (indicating a shareable resource)
- Use bus shared (indicating that a partition is sharing a bus owned by another partition)

Accordingly, there are two choices available when partitioning a machine:

- Bus level partitioning, where a bus along with all the IOPs attached to it is dedicated to a single partition. In this case, there is no sharing of IOP resources.
- IOP level partitioning, where a bus is shared among partitions and the IOPs attached to it can be independently switched from partition to partition as the need arises.

A.14 Addition and Removal

Addition and *removal* are the terms used to switch resources between partitions. When resources are removed from a logical partition, they become available for addition to another logical partition.

Some addition and removal actions may require an entire system restart. Other such actions may only require a secondary logical partition restart.

A.15 OptiConnect

OptiConnect is a function that allows you to connect multiple systems using a physical high-speed, fiber-optic bus. Virtual OptiConnect emulates the physical OptiConnect hardware to achieve the same objective between logical partitions.

To use Virtual Opticonnect, all you need is to purchase the software license for OptiConnect or OptiMover and configure your logical partitions to enable the Virtual OptiConnect function.

In addition to Virtual OptiConnect, inter-partition communication can also be achieved with standard LAN/WAN facilities. In such a case, each partition needs to have a dedicated communication adapter.

Appendix B. Sample RPG and C Programs

The following examples show how to use the unblocked MI instruction, MATMATRI, to retrieve partition resource information.

B.1 Sample ILE RPG Program — Output to Spooled File

```
0001.00 H*************************************************************************
0002.00 H* COMPILE THIS PROGRAM WITH DFTACTGRP(*NO)                               *
0003.00 H*                                                                        *
0004.00 H* Sample ILE RPG program using the MI instruction MATMATR to             *
0005.00 H* materialize partition information on the AS/400.                       *
0006.00 H*************************************************************************
0007.00 F*
0008.00 F* Print file specifications for output
0009.00 F*
0010.00 FQSYSPRT   O    F  132        PRINTER
0011.00 F*
0012.00 D*
0013.00 D* Prototype of MI MATMATR instruction
0014.00 D*
0015.00 DMatMAtr          PR                  EXTPROC('_MATMATR1')
0016.00 D                                  *   VALUE
0017.00 D                                  2   CONST
0018.00 D*
0019.00 D* Working variables for Materialize
0020.00 D*
0021.00 DAttribPtr        s                 *   inz(%addr(Attributes))
0022.00 DMatOption        s                 2   inz(x'01E0')
0023.00 D*
0024.00 D* Receiver variable for Materialize
0025.00 D*
0026.00 DAttributes       DS              512
0027.00 D BytPrv                           10i 0
0028.00 D BytAvl                           10i 0
0029.00 D NumParts                          3u 0
0030.00 D CurPart                           3u 0
0031.00 D PriPart                           3u 0
0032.00 D Reserved                          5
0033.00 D LogSerial                        10
0034.00 D Reserved2                         6
0035.00 D MinProc                           5u 0
0036.00 D MaxProc                           5u 0
0037.00 D CurProc                           5u 0
0038.00 D Reserved3                         2
0039.00 D CfgMinMem                        10u 0
0040.00 D CfgMaxMem                        10u 0
0041.00 D CurAvlMem                        10u 0
0042.00 D MinIntPerf                        3u 0
0043.00 D MaxIntPerf                        3u 0
0044.00 D CurIntPerf                        3u 0
0045.00 D Reserved4                         1
0046.00 D*
0047.00 D Description     S                 25   inz(' ')
0048.00 D ValueN          S                 10  0 inz(0)
0049.00 D ValueC          S                 10   inz(' ')
0050.00 C*
0051.00 C* Set Bytes Provided to size of Recevier Variable (Attributes
0052.00 C*
```

```
0053.00 C                   eval      BytPrv = %size(Attributes)
0054.00 C*
0055.00 C* Use MATMATR MI instruction
0056.00 C*
0057.00 C                   callp     MatMAtr(AttribPtr: MatOption)
0058.00 C*
0059.00 C* Determine if information returned
0060.00 C*
0061.00 C                   if        BytAvl >= 56
0062.00 C*
0063.00 C* Read though the detail and output it to the printer
0064.00 C*
0065.00 C                   Except    Header
0066.00 C*
0067.00 C* Total number of partitions that have been created
0068.00 C*
0069.00 C                   eval      Description = 'Total Partitions Created'
0070.00 C                   eval      ValueN = NumParts
0071.00 C                   Except    DetailN
0072.00 C*
0073.00 C* Current partition
0074.00 C*
0075.00 C                   eval      Description = 'Current Partition Number'
0076.00 C                   eval      ValueN = CurPart
0077.00 C                   Except    DetailN
0078.00 C*
0079.00 C* Logical Serial Number
0080.00 C*
0081.00 C                   eval      Description = 'Logical Serial Number   '
0082.00 C                   eval      ValueC = LogSerial
0083.00 C                   Except    DetailC
0084.00 C*
0085.00 C* Minimum No Processors
0086.00 C*
0087.00 C                   eval      Description = 'Minimum No Processors   '
0088.00 C                   eval      ValueN = MinProc
0089.00 C                   Except    DetailN
0090.00 C*
0091.00 C* Maximum No Processors
0092.00 C*
0093.00 C                   eval      Description = 'Maximum No Processors   '
0094.00 C                   eval      ValueN = MaxProc
0095.00 C                   Except    DetailN
0096.00 C*
0097.00 C* Current No Processor
0098.00 C*
0099.00 C                   eval      Description = 'Current No Processors   '
0100.00 C                   eval      ValueN = CurProc
0101.00 C                   Except    DetailN
0102.00 C*
0103.00 C* Minimum Amount Memory
0104.00 C*
0105.00 C                   eval      Description = 'Minimum Amount Memory   '
0106.00 C                   eval      ValueN = CfgMinMem
0107.00 C                   Except    DetailN
0108.00 C*
0109.00 C* Maximum Amount Memory
0110.00 C*
0111.00 C                   eval      Description = 'Maximum Amount Memory   '
0112.00 C                   eval      ValueN = CfgMaxMem
0113.00 C                   Except    DetailN
0114.00 C*
0115.00 C* Current Amount Memory
```

```
0116.00 C*
0117.00 C                   eval      Description = 'Current Amount Memory   '
0118.00 C                   eval      ValueN = CurAvlMem
0119.00 C                   Except    DetailN
0120.00 C*
0121.00 C* Minimum Interactive Performance %
0122.00 C*
0123.00 C                   eval      Description = 'Minimum Interactive Perf'
0124.00 C                   eval      ValueN = MinIntPerf
0125.00 C                   Except    DetailN
0126.00 C*
0127.00 C* Maximum Interactive Performance %
0128.00 C*
0129.00 C                   eval      Description = 'Maximum Interactive Perf'
0130.00 C                   eval      ValueN = MaxIntPerf
0131.00 C                   Except    DetailN
0132.00 C*
0133.00 C* Current Interactive Performance %
0134.00 C*
0135.00 C                   eval      Description = 'Current Interactive Perf'
0136.00 C                   eval      ValueN = CurIntPerf
0137.00 C                   Except    DetailN
0138.00 C                   Except    End
0139.00 C*
0140.00 C                   else
0141.00 C                   eval      ValueC = ' '
0142.00 C                   eval      Description = '****Error Occurred*******'
0143.00 C                   Except    DetailC
0144.00 C                   Except    End
0145.00 C                   endif
0146.00 C*
0147.00 C* exit sample program
0148.00 C*
0149.00 C                   eval      *inlr = '1'
0150.00 C                   return
0151.00 C*
0152.00 OQSYSPRT   E  1         HEADER         2 03
0153.00 O                                           24 'Partition Information'
0154.00 O          E            DetailN     1
0155.00 O                       Description         28
0156.00 O                       ValueN        3     45
0157.00 O          E            DetailC     1
0158.00 O                       Description         28
0159.00 O                       ValueC              47
0160.00 O          E            END         2
0161.00 O                                           28 '*** END OF REPORT ***'
```

B.2 Sample C Program — Output to Screen

```
#include <stdio.h>
#include <QSYSINC/MIH/MATMATR>

_MMTR_Template_T  machine_attributes;

int main(void)

{
machine_attributes.Options.Template_Size = sizeof(_MMTR_Template_T);
     matmatr ( &machine_attributes, _MMTR_LPAR_INFO) ;

printf ("Total Partitions Created: %d \n",
```

```
     machine_attributes.Options.Data.LparInfo.num_of_partition);
printf ("Current Partition Number: %d \n",
     machine_attributes.Options.Data.LparInfo.partition_id);
printf ("Primary Partition number: %d \n",
     machine_attributes.Options.Data.LparInfo.prim_partition);
printf ("Logical Serial Number   : %s \n",
     machine_attributes.Options.Data.LparInfo.logical_ser_num);
printf ("Minimum No Processors   : %i \n",
     machine_attributes.Options.Data.LparInfo.min_processors);
printf ("Maximum No Processors   : %i \n",
     machine_attributes.Options.Data.LparInfo.max_processors);
printf ("Current No Processors   : %i \n",
     machine_attributes.Options.Data.LparInfo.cur_processors);
printf ("Minimum Amount Memory   : %i \n",
     machine_attributes.Options.Data.LparInfo.cur_min_memory);
printf ("Maximum Amount Memory   : %i \n",
     machine_attributes.Options.Data.LparInfo.cur_max_memory);
printf ("Current Amount Memory   : %i \n",
     machine_attributes.Options.Data.LparInfo.cur_avail_memory);
printf ("Minimum Interactive Perf: %i \n",
     machine_attributes.Options.Data.LparInfo.min_percentage_int_work);
printf ("Maximum Interactive Perf: %i \n",
     machine_attributes.Options.Data.LparInfo.max_percentage_int_work);
printf ("Current Interactive Perf: %i \n",
     machine_attributes.Options.Data.LparInfo.cur_percentage_int_work);
}
```

Appendix C. Accessing LPAR Functions

Table 17 explains the logical partitioning menu options available from the DST and SST screens in both the primary and secondary partitions. The initial logical partitioning menu can be reached by selecting option 11 from the main DST menu or option 5 from the main SST menu.

Table 17. Logical Partitions Menu Options

Logical Partitioning Function	**Primary Partition**		**Secondary Partition**	
	DST	SST	DST	SST
Option 1 - Display partition information	Y	Y	Y	Y
1. Display partition status	Y	Y	Y(1)	Y(1)
2. Display partition processing configuration	Y	Y	Y(1)	Y(1)
3. Display allocated resources	Y	Y	Y(1)	Y(1)
4. Display available resources	Y	Y	Y	Y
5. Display system I/O resources	Y	Y	N	N
6. Display partition release level	Y	Y	N	N
7. Display secondary partition reference code history	Y	Y	N	N
Option 2 - Work with partition status	Y	Y	Y(1)	Y(1)
Option 3 - Work with partition configuration	Y	Y(2)	N	N
Option 4 - Recover configuration data	Y	Y	Y	Y(3)
1. Recover primary partition configuration data	Y	N	N	N
2. Update configuration data	Y	Y	N	N
3. Clear non-configured disk unit configuration data	Y	N	Y(1)	N
4. Clear non-reporting logical partitioning resources	Y	N	N	N
5. Accept load source disk unit	Y	N	Y(1)	N
6. Copy configuration data to other side	Y	N	Y(1)	N

Logical Partitioning Function	Primary Partition		Secondary Partition	
7.Clear configuration data	Y	N	N	N
Option 5 - Create a new partition	Y	N	N	N

Notes:
Options with an "N" in the SST column for the primary partition are displayed. If they are selected, an error message is sent.

Options with an "N" in either the DST or SST columns of the secondary partitions are not displayed with the exception of note 3 (below).

1. You can only work with or display information related to this secondary partition.
2. Options 1, 2, and 10 are displayed but can only be selected when in DST of the primary partition.
3. This menu option is displayed, but an error message is sent if selected.

Security Level Required to Access the Menu Options

You must use either the `QSECOFR` or `22222222` ID to access the menu options from within DST. To access the menu options in SST, the user profile that you use must have **SERVICE* specified as one of its special authorities.

Appendix D. Operations Console Connection for Partitioned AS/400

There are four types of console devices commonly found in an AS/400 system. These devices are:

- Twinaxial workstation
- ASCII workstation
- Client Access console on PC
- Operations Console on PC

However, logical partitions on the AS/400 system only support the twinaxial console and the Operations Console. This appendix offers tips and hints on how to set up Operations Console in an AS/400 partition.

D.1 What Operations Console Is

The Operations Console feature comes standard with Client Access for Windows 95, 98, and NT V3R2M0 or Client Access Express for Windows V4R4M0. It also requires OS/400 V4R3M0 or later release on the AS/400 system.

AS/400 Operations Console allows you to use a PC with Windows 95 or Windows NT to access and control, either remotely or locally, the AS/400 console and control panel. Each AS/400 partition must have a console attached to it.

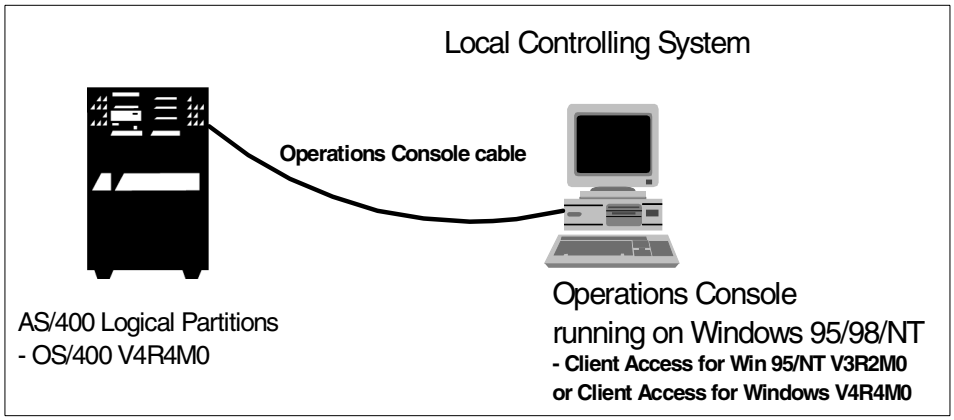


Figure 85. Operations Console Attachment

The Operations Console PC is directly connected to an AS/400 system through a special Operations Console cable (p/n 97H7557, feature #0367 for

#2721 attachment, p/n 97H7556, feature #0328 for #2699 attachment and #97H7555 for #2609 or #2612 attachment). It is connected between the PC serial port and the AS/400 communication adapter (port 0 if a two-line communication adapter). This PC is called *Local Controlling System (LCS)*. The LCS can be configured to access the AS/400 console function on either the primary partition or the secondary partition. The LCS can also be configured to access the AS/400 control panel function on the *primary partition*. In this case, the PC must also be connected to the AS/400 control panel port through a special Panel cable (p/n 97H7584, feature #0380 for SPD bus system unit and p/n 97H7591, feature #0381 for PCI bus system unit) and a second serial port on the PC. There is no control panel support on the LCS if it is attached to the secondary partition. Only one LCS can be connected to one AS/400 partition.

The LCS does not use the previous SNA-over-Async type of communication, like the Client Access console, between the AS/400 system and the LCS. Instead, the LCS acts as a router between the AS/400 system and the Personal Communication 5250 emulation using TCP/IP through the TCP/IP loop back function on the PC. The 5250 emulation for the AS/400 console function can be run on another PC through a dial-up connection to the LCS. This remote PC is called *Remote Controlling System (RSC)*. To use RSC for remote access, the LCS must be a Windows NT (workstation) PC.

D.2 Direct-Attached Operations Console in Secondary Partition

To attach the Operations Console to an AS/400 partition, either the primary partition or a secondary partition, you need to perform the following steps. You can refer to *Operations Console Setup,* SC41-5508, for the details.

Tips

Please check the following items before the Operations Console is started:

1. Ensure that the Operations Console cable is attached correctly.
2. Install the AS/400 Operations Console Connection null modem in your Windows 95, 98, or NT PC.
3. Select the console IOP as the default ECS resource in partition configuration.
4. Select the Operations Console as the default console type in the DST.
5. If you have twinaxial IOA attached to the same console IOP, you need to switch off the attached twinaxial workstation with the addresses port 0 address 0, port 0 address 1 and port 1 address 0 in order to use the Operations Console as the system console.

The following list highlights the areas that you need to pay attention to in LPAR environment. Complete these steps:

1. Connect the Operations Console cable correctly.

   Normally, port 0 is the left-hand port of feature code #2721 when it is attached to system bus 1 of 620/S20/720. However, it will be the right-hand port of feature code #2721 if the #2721 is attached to system bus 2 of Model 620/S20/720. It is because the #2721 on the system bus 2 is turned over when compared to the position on the system bus 1.

   Port 0 is the lower port for the communication IOA, such as #2699, for the SPD card.

2. Install Client Access for Windows 95/NT V3R2 or the Client Access Express for Windows V4R4 with Operations Console on your PC.

   You need to configure the AS/400 Operations Console null modem (see Figure 86 on 180).

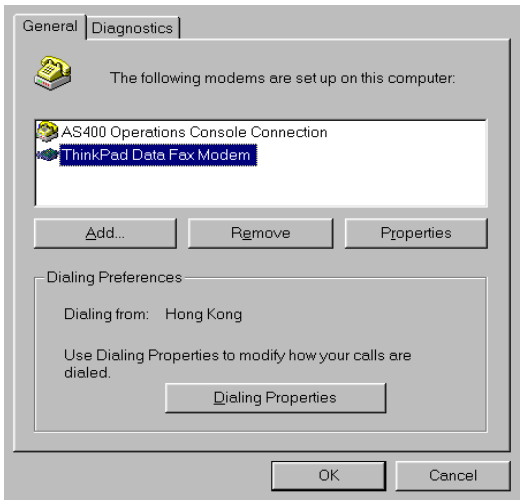


Figure 86. AS/400 Operations Console Modem

If you have not installed the null modem, you can install the AS/400 Operations Console null modem from `X:\Program Files\IBM\Client Access\Aoc\inf`. Here, `X` is the directory where you installed the Client Access base code.

3. Identify the communication IOP or the combined function IOP or the MFIOP that has the Operations Console attached:
   - #2809 (combined function IOP with #2721 attached) for PCI bus system
   - #2629 (communication IOP with #2699 attached) for SPD bus
   - #2623 (communication IOP with #2609 attached) for SPD bus
   - MFIOP
4. Select (or tag) the console IOP in the partition configuration Figure 87 on 181).

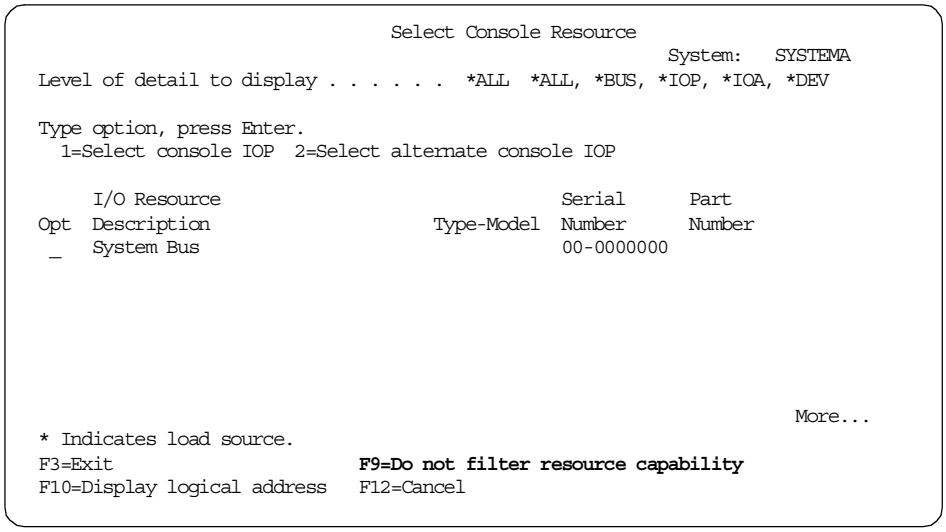


Figure 87. Select Console IOP with Filtering

The communication IOP, combined function IOP, or MFIOP to which the Operations Console is attached, is not shown as seen in Figure 87 on 181. The IOP is filtered out by the built-in filter rules since it has no twinaxial workstation IOA attached. You need to press **F9** (Do not filter resource capability) to get back the communication IOP, combined function IOP, or the MFIOP for your selection as shown in Figure 88 on 182.

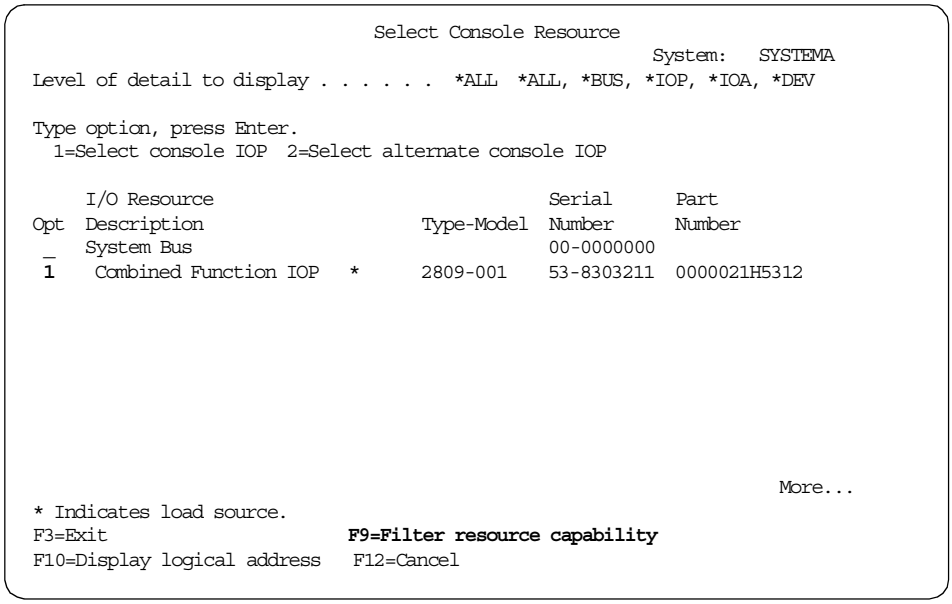


Figure 88. Select Console IOP without the Filtering

Then, select the IOP with the Operations Console PC attached.

5. Select the console IOP as the Electronic Customer Support IOP also.

   To have the AS/400 partition correctly pick up the communication adapter, you must select the console IOP in the previous step as your default Electronic Customer Support Resource.

   Go to the Work with Partition Configuration display (Figure 89 on 183), and enter option 9 next to the secondary partition, for example, LPAR2.

```
                    Work with Partition Configuration
                                                              System:   SYSTEMA
Type option, press Enter.
  7=Select console resource  8=Select alternate IPL resource
  9=Select default electronic customer support resource
 10=Delete partition

        Partition
Option  Identifier  Name
  __        0       PRIMARY
  9_        1       LPAR2









F3=Exit   F11=Work with partition status   F12=Cancel   F23=More options
```

Figure 89. Work with Partition Configuration Menu

The < sign next to the Combined Function IOP in Figure 90 indicates that the IOP has already been selected as the console IOP.

```
           Select Default Electronic Customer Support Resource
                                                              System:   RCHASM25
Level of detail to display . . . . . *ALL  *ALL, *BUS, *IOP, *IOA, *DEV

Type option, press Enter.
  1=Select IOP

     I/O Resource                              Serial      Part
Opt  Description                 Type-Model   Number      Number
  _   System Bus                               00-0000000
  1    Combined Function IOP   *<   2809-001   53-8303211  0000021H5312









                                                                      More...
* Indicates load source.
F3=Exit                     F9=Do not filter resource capability
F10=Display logical address   F12=Cancel
```

Figure 90. Select Default Electronic Customer Support Resource Menu

Select the console IOP as the default Electronic Customer Support IOP.

6. Configure the Operations Console on your PC for the connection.

   Refer to Chapter 37 of *Client Access for Windows 95/NT Setup V3R2,* SC41-3512, for the details.

7. Start the Operations Console and power on the secondary partition.
8. When the Operations Console DST window pops up on the PC, use any one of the DST user IDs to sign on:

   - QSECOFR — Security capability
   - 22222222 — Full capability
   - 11111111 — Basic capability

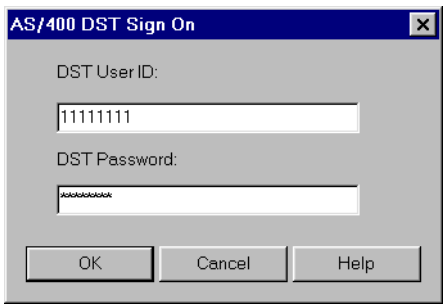


Figure 91. Operations Console DST Sign On Window

   Use the basic user ID `11111111` to sign on to the Operations Console.

9. If this is a first time installation, continue with the disk formatting and SLIC installation.
10. Select **Operations Console** as your console device in DST.

    You should now be at the IPL or Install the System menu as shown in Figure 92 on 185.

    a. Select option **3** (Use Dedicated Service Tools (DST)) and sign on as `QSECOFR`.

```
                    IPL or Install the System
                                                          System:   LPAR2
Select one of the following:

     1. Perform an IPL
     2. Install the operating system
     3. Use Dedicated Service Tools (DST)
     4. Perform automatic installation of the operating system
     5. Save Licensed Internal Code




Selection
     3

Licensed Internal Code - Property of IBM 5769-999 Licensed
Internal Code (c) Copyright IBM Corp.  1980, 1999.  All
rights reserved.  US Government Users Restricted Rights -
Use duplication or disclosure restricted by GSA ADP schedule
Contract with IBM Corp.
```

Figure 92. IPL or Install the System Menu

b. Select option **5** (Work with DST environment). See Figure 93.

```
                Use Dedicated Service Tools (DST)
                                                      System:   LPAR2
Select one of the following:

     1. Perform an IPL
     2. Install the operating system
     3. Work with Licensed Internal Code
     4. Work with disk units
     5. Work with DST environment
     6. Select DST console mode
     7. Start a service tool
     8. Perform automatic installation of the operating system

    10. Work with remote service support
    11. Work with system partitions



Selection
     5

F3=Exit   F12=Cancel
```

Figure 93. DST Main Menu

c. Select option **2** (System devices). See Figure 94 on 186.

```
                        Work with DST Environment
                                                              System:   LPAR2
Select one of the following:

     1. Active service tools
     2. System devices
     3. DST user profiles
     4. System values




Selection
     2

F3=Exit   F12=Cancel
```

Figure 94. Work with DST Environment Menu

d. Select option **6** (Console mode). See Figure 95.

```
                        Work with System Devices
                                                              System:   LPAR2
Select one of the following:

     1. Printers
     2. Tape devices
     3. Diskette devices
     4. Optical devices
     5. Alternate installation device
     6. Console mode




Selection
     6

F3=Exit   F12=Cancel
```

Figure 95. Work with System Devices Menu

e. Select option **2** (Operations Console) and press **Enter**.

```
                        Select Console Type
                                                              System:   LPAR2
Select one of the following:

     1. Local Console
     2. Operations Console




Selection
     2

F12=Cancel
```

Figure 96. Select Console Type Menu

11.Exit from DST and continue the system installation.

Appendix E. Interactive and Batch CPW Determination

If you need to consolidate a number of AS/400 systems into one AS/400 partition, you must evaluate both the interactive and batch performance requirements for the combined workload. The methodology for the performance determination or capacity planning is no different from the consolidation of two AS/400 systems into one system.

This appendix provides the steps to evaluate both interactive and batch performance requirements of the combined workload using the BEST/1 tool.

E.1 Overview of the Methodology

In general, you can characterize the system workload in many different ways. For example, you can characterize the system workload according to the business function that the jobs perform. You may have Order Entry, Accounts Receivable, or Accounts Payable workloads during the day time, and Customer Master Update and General Ledger Update workloads at night. Or, you can simply characterize the system loading into interactive workload during the day time and batch workload at night.

Use the following steps to determine both the interactive and the batch performance requirements:

1. Define or characterize the workloads.
2. Define the performance expectation. It can be interactive response time, batch runtime or the number of orders processed.
3. Collect the performance data of the workloads using the AS/400 performance monitor, `STRPFRMON` command.
4. Analyze the performance data and create the BEST/1 model for the workloads. Workload can be classified according to the user ID, subsystem name, job name, job type, or job number. You can refer to the BEST/1 manual or Chapter 4 of the redbook *AS/400 Server Capacity Planning,* SG24-2159, for details.
5. Compare the measured result with the customer's performance expectation.
6. Combine the workloads if they run concurrently and project the machine requirement in terms of the CPU usage, main storage size, the expected performance, and the partition configuration.
7. Derive the interactive CPW number and batch CPW number.

E.2 Interactive Performance Determination

You should have performed step 1 to step 5 in E.1, “Overview of the Methodology” on page 189, before you can continue to determine the interactive CPW requirement. Now you have a clear idea of your current performance level and whether the current performance level matches the expectation. Use the following steps to combine the interactive workloads and evaluate them on a user defined AS/400 partition model in BEST/1:

1. Create a user-defined model based on the target AS/400 system and the proposed LPAR configuration. You can refer to 6.7, “Creating a New Hardware Feature in BEST/1” on page 99, for details.
2. If you want to combine two or more workloads into one AS/400 partition, you need to save the workload in the BEST/1 model that you created.

   For example, you created two BEST/1 models, INTER01 in library QGPL and M53SINT in library QPFR_TIM53, for the interactive workloads of two different AS/400 systems that you want to consolidate.

   The following steps help you to combine the workloads into the base model. The base model is one of the BEST/1 models that you created for the consolidation. It is the starting base for the system performance projection or evaluation. It is normally the biggest system or the one with most of the communication lines and LAN adapters among the AS/400 systems. Let us assume the base model is M53SINT in library QPFR_TIM53. You need to save the workloads from INTER01 and add them to the model M53SINT.

   a. Start AS/400 Performance Tools and work with the BEST/1 model that you want to save the workload.
   b. Choose option **5** to work with the model INTER01.

```
                          Work with BEST/1 Models

Library . . . . .   QGPL        Name

Type options, press Enter.
  1=Create   3=Copy   4=Delete   5=Work with   6=Print   7=Rename

Opt  Model      Text                                      Date      Time

 5   INTER01    Interactive Performance                   02/04/99  14:51:30




                                                                        Bottom
Command
===>
F3=Exit   F4=Prompt   F5=Refresh        F9=Retrieve   F12=Cancel
F15=Sort by model     F16=Sort by text   F19=Sort by date and time
```

Figure 97. Choosing Option 5 to Work with the Model INTER01

c. Save the workloads of the current model.

d. Choose option **1** (Work with workloads) to work with the workloads of the model INTER01.

```
                          Work with BEST/1 Model

Performance data . . . :   QPFRDATA (OCTEST1)
Model/Text . . . . . . :   INTER01     Interactive Performance


Select one of the following:

     1. Work with workloads
     2. Specify objectives and active jobs

     5. Analyze current model
     6. Analyze current model and give recommendations
     7. Specify workload growth and analyze model

    10. Configuration menu
    11. Work with results

                                                                        More...
Selection or command
===>  1
F20=More BEST/1 options   F21=Basic user level   F24=More keys
```

Figure 98. Choosing Option 1 to Work with the Workloads of the Model INTER01

e. Save the workloads to the base model. The base model is normally the largest system among the systems that you want to consolidate.

f. Choose option **8** to select the required workloads.

```
                          Work with Workloads

Model/Text:   INTER01     Interactive Performance

Type options, press Enter.
  1=Create   2=Change   3=Copy   4=Delete   5=Display   6=Print   7=Rename
  8=Save workload to workload member       9=Edit transactions

Opt     Workload       Text

 8      CLIENTAC4      Interactive performance data from SG system
 8      INTERACTIV     Interactive performance data from SG system
 8      NONINTER       Interactive performance data from SG system
 8      QDEFAULT       Interactive performance data from SG system









                                                                      Bottom
F3=Exit   F6=Add saved workload   F9=Add predefined workload   F12=Cancel
F13=Combine workloads
```

Figure 99. Choosing Option 8 to Select the Required Workloads

g. Save the workload to the base model.

   Change the name of the workload to a name with system identification and save it to the library of the base model.

   For example, you can rename the CLIENTAC4 workload as `SYSACA4` and fill in QPFR_TIM53 for the base model's library (see Figure 100 on page 193).

```
                    Save Workload to Workload Member

Change values if desired, press Enter.
  Member . . . . . .   SYSACA4        Name
    Library  . . . . .    QPFR_TIM53   Name

  Text . . . . . .   Interactive performance data from SG system

  Replace  . . . . . .   N             Y=Yes, N=No

  CPU architecture . . . . . .   *RISC         *CISC, *RISC









F12=Cancel
```

Figure 100. Save Workload to Workload Member

After you save all of the required workloads, you return to the Work with Workloads display.

```
                            Work with Workloads

Model/Text:   INTER01     Interactive Performance

Type options, press Enter.
  1=Create   2=Change   3=Copy   4=Delete   5=Display   6=Print   7=Rename
  8=Save workload to workload member       9=Edit transactions

Opt     Workload       Text

        CLIENTAC4      Interactive performance data from SG system
        INTERACTIV     Interactive performance data from SG system
        NONINTER       Interactive performance data from SG system
        QDEFAULT       Interactive performance data from SG system









                                                                      Bottom
F3=Exit   F6=Add saved workload   F9=Add predefined workload   F12=Cancel
F13=Combine workloads
```

Figure 101. Work with Workloads

Press **F12** to exit.

h. Now, you are ready to import the saved workloads to the base model.

Go back to Work with BEST/1 Model and work with the base model, M53SINT in library QPFR_TIM53.

Choose option **1** (Work with workloads) to work with the base model's workloads.

```
                          Work with BEST/1 Model

Performance data . . . :   QPFR_TIM53 (Q983510730)
Model/Text . . . . . . :   M53SINT


Select one of the following:

     1. Work with workloads
     2. Specify objectives and active jobs

     5. Analyze current model
     6. Analyze current model and give recommendations
     7. Specify workload growth and analyze model

    10. Configuration menu
    11. Work with results

                                                                      More...
Selection or command
===> 1
F20=More BEST/1 options   F21=Basic user level   F24=More keys
```

Figure 102. Work with Best/1

i. Press **F6** to add the saved workload to the base model (Figure 103 on page 195).

```
                          Work with Workloads

Model/Text:   M53SINT

Type options, press Enter.
  1=Create   2=Change   3=Copy   4=Delete   5=Display   6=Print   7=Rename
  8=Save workload to workload member       9=Edit transactions

Opt     Workload       Text

        CLIENTAC4      Interactive performance data from HK system
        INTERACTIV     Interactive performance data from HK system
        NONINTER       Interactive performance data from HK system
        QDEFAULT       Interactive performance data from HK system



                                                                      Bottom
F3=Exit   F6=Add saved workload   F9=Add predefined workload   F12=Cancel
F13=Combine workloads
```

Figure 103. Work with Workloads

j. Select the previous saved workload. Enter a `1` next to the required workloads for selection and press **Enter**. You need to select the workload one by one. See Figure 104.

```
                          Add Saved Workload

Library . . . . .   QPFR_TIM53   Name

Type option, press Enter.
  1=Select

Opt  Member      Text                                    Date      Time
 1   SYSANINTR   Interactive performance data from SG     02/04/99  15:45:28
     SYSAINTR    Interactive performance data from SG     02/04/99  15:45:19
     SYSADFT     Interactive performance data from SG     02/04/99  15:45:13
     SYSACA4     Interactive performance data from SG     02/04/99  15:45:05









                                                                      Bottom
F3=Exit   F5=Refresh   F12=Cancel   F15=Sort by member   F16=Sort by text
F19=Sort by date and time
```

Figure 104. Add Saved Workload

The workloads are added successfully.

```
                              Work with Workloads

Model/Text:   M53S

Type options, press Enter.
  1=Create   2=Change   3=Copy   4=Delete   5=Display   6=Print   7=Rename
  8=Save workload to workload member       9=Edit transactions

Opt     Workload       Text

        CLIENTAC4      Interactive performance data from HK system
        INTERACTIV     Interactive performance data from HK system
        NONINTER       Interactive performance data from HK system
        QDEFAULT       Interactive performance data from HK system
        SYSACA4        Interactive performance data from SG system
        SYSADFT        Interactive performance data from SG system
        SYSAINTR       Interactive performance data from SG system
        SYSANINTR      Interactive performance data from SG system


                                                                      Bottom
F3=Exit   F6=Add saved workload   F9=Add predefined workload   F12=Cancel
F13=Combine workloads
Workload SYSACA4 has been read.  Check ASP, pool, and comm line assignments
```

Figure 105. Work with Workload

k. You may need to add additional storage pools, disk units to ASPs, or communication adapters to the base model to support the combined workload.

l. You can continue with the normal performance projection by specifying the growth objective and choosing the user-defined AS/400 partition model.

   Refer to 6.7, “Creating a New Hardware Feature in BEST/1” on page 99, to learn how to create a user-defined AS/400 model for an AS/400 partition.

Once you satisfy the projected performance on the user-defined AS/400 partition model, you can convert the Normal Performance rating of the user-defined AS/400 partition model back to an interactive CPW number.

```
                              Work with CPU Models

Type option, press Enter.
  1=Create   2=Change   3=Copy   4=Delete   7=Rename

                  System   Archi-  --Relative Perf---   Number of  Max Stor
Opt  CPU Model     Unit   tecture    Normal    Server  Processors    (MB)

           B10    9404    *CISC       1.00                 1          16
     650   LP04   9406    *RISC     262.35                 4       32768
     730   2A6A   9406    *RISC      18.10    147.36        1       24576
     730   2A6B   9406    *RISC      29.33    147.36        1       24576
     730   2A6C   9406    *RISC      59.02    147.36        1       24576
     730   2A6D   9406    *RISC     147.36                 1       24576
     730   2A6E   9406    *RISC      17.74    274.80        2       24576
     730   2A6F   9406    *RISC      29.33    274.80        2       24576
     730   2B6A   9406    *RISC      58.29    274.80        2       24576
     730   2B6B   9406    *RISC     145.91    274.80        2       24576
     730   2B6C   9406    *RISC     274.80                 2       24576
     730   2B6D   9406    *RISC      17.02    505.43        4       24576
     730   2B6E   9406    *RISC      54.67    505.43        4       24576
                                                                    More...
F3=Exit   F9=Display equivalent BEST/1 and IBM CPU models   F12=Cancel
```

Figure 106. User-Defined AS/400 Partition Model — 650-LP04

For example, the projected performance fits into a user-defined AS/400 partition Model 640-LP04 with interactive CPU utilization around 70% and interactive response time less than 1.5 seconds. The user-defined AS/400 partition model 640-LP04 is created by using the Model 9406-640 #2240 8-way processor with a relative performance rating of 524.70 as the base. However, the Model 640-LP04 has only four processors and a relative performance rating of 262.35. Since the interactive CPW of a 9406-640 #2240 processor is 1,794, the required interactive for the combined workload is:

```
1,794 x (262.35 / 524.70) = 897
```

E.3 Batch Performance Determination

You should have performed step 1 through step 5 in E.1, “Overview of the Methodology” on page 189, before you continue to determine the batch performance requirement. Now you have a clear idea of your current performance level and whether the current performance level matches the expectation.

For example, a customer has two AS/400 servers running the same ERP application. One is located in Hong Kong and the other is located in

Singapore. Both AS/400 servers have the same capacity, 9406-53S #2157 processor with 2 GB of main storage. The system loading for both places is roughly the same for the batch processing at night. The batch processing is normally started at 7:30pm and ended at 11:30pm. Figure 107 shows the CPU utilization versus time chart of a typical batch processing of the customer.

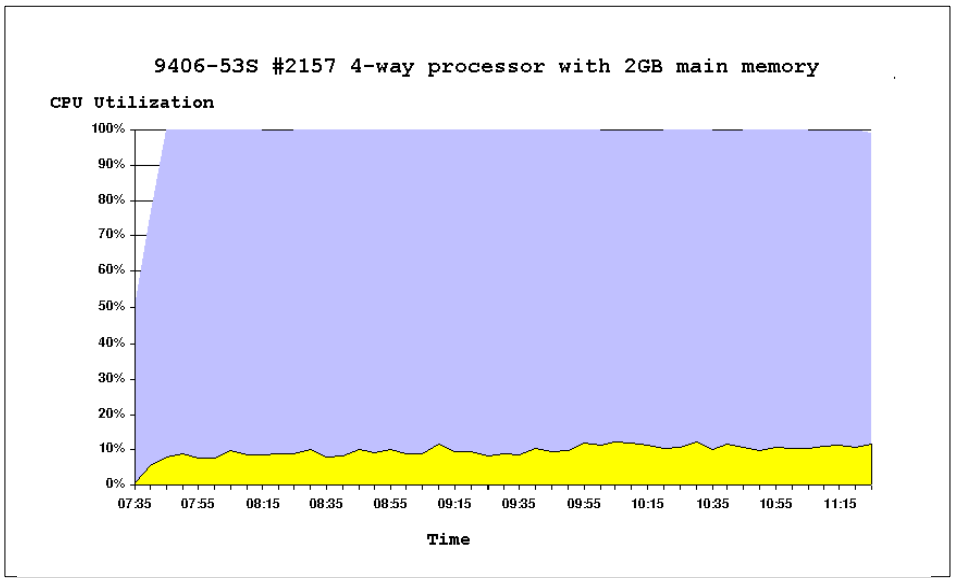


Figure 107. CPU Utilization versus Time Chart

The customer wants to consolidate the two 9406-53S servers into one of the two partitions of a new Model 9406-S40 #2207 8-way processor. The other partition is used for the new workflow application running on Domino/400. What is the required batch CPW for the combined workload?

Let us look at one system first. Figure 108 on page 199 shows that the 9406-53S has been driven up to 98.4% and the disk arm utilization is low, only 2%. Therefore, disk arm contention is not the bottleneck. To complete the batch processing within four hours, you need an AS/400 server with comparable batch performance. The client/server CPW (or batch CPW) of the 9406-53S #2157 processor is 650. Therefore, you need a partition with at least 1,300 (650 x 2) batch CPW.

```
                          Display Analysis Summary

                          Stor     CPU  -Disk IOPs--  -Disk Ctls--  -Disk Arms--
Period    CPU Model       (MB)    Util   Nbr    Util   Nbr    Util   Nbr    Util
Base      53S 2157        2048    98.4     5     5.6    11      .8    70     2.2
+10%      53S 2157        2048    99.0     5     5.8    11      .9    70     2.3
+10%      53S 2157        2048    99.0     5     6.2    11      .9    70     2.5
+10%      53S 2157        2048    99.0     5     6.7    11     1.0    70     2.7
+10%      53S 2157        2048    99.0     5     7.3    11     1.1    70     3.0

                                                                          Bottom
          ----Inter Rsp Time----  -------Inter--------  -----Non-Inter------
Period     Local    LAN     WAN   CPU Util    Trans/Hr  CPU Util    Trans/Hr
Base          .5     4.1      .0       .0          47      98.4     14032018
+10%          .5     4.1      .0       .0          52      99.0     13898859
+10%          .5     4.1      .0       .0          57      99.0     13629150
+10%          .5     4.1      .0       .1          63      98.9     13332477
+10%          .5     4.1      .0       .1          69      98.9     13006129

                                                                          Bottom
F3=Exit   F10=Re-analyze   F11=Alternative view          F12=Cancel
F15=Configuration menu     F17=Analyze multiple points   F24=More keys
```

Figure 108. BEST/1 Projection for the Batch Processing

To project the performance requirement, you need to save the workloads from the BEST/1 model for one of the Model 9406-53S servers. Please refer to E.2, “Interactive Performance Determination” on page 190, for details.

We assume that you saved workloads to the base model, M53S and added to the workload profile as shown in Figure 109 on page 200.

```
                          Work with Workloads

Model/Text:   M53S

Type options, press Enter.
  1=Create   2=Change   3=Copy   4=Delete   5=Display   6=Print   7=Rename
  8=Save workload to workload member       9=Edit transactions

Opt     Workload       Text

        CLIENTAC4      Batch performance data for HK system
        INTERACTIV     Batch performance data for HK system
        NONINTER       Batch performance data for HK system
        QDEFAULT       Batch performance data for HK system
        SYSBCA4        Batch performance data for SG system
        SYSBDFT        Batch performance data for SG system
        SYSBINTR       Batch performance data for SG system
        SYSBNINTR      Batch performance data for SG system


                                                                      Bottom
F3=Exit   F6=Add saved workload   F9=Add predefined workload   F12=Cancel
F13=Combine workloads
```

Figure 109. Workload Profile of M53S

Now, go to Work with BEST/1 Model display and select option **10** (Configuration menu).

As shown in Figure 110 on page 201, select option **1** (Change CPU and other resource values) and press **Enter**.

```
                                Configuration

CPU Model  . . . . . :   2157      Comm IOPs  . . . . . :   1
Main stor (MB) . . . . . :   2048      LAN lines  . . . . . :   1
Main stor pools  . . . . . :   3       WAN lines  . . . . . :   0
Disk IOPs  . . . . . :   5
Disk ctls  . . . . . :   11        Multifunction IOPs . . . . :   1
Disk arms  . . . . . :   70        Disk IOAs  . . . . . :   0
ASPs . . . . . :   1         Comm IOAs  . . . . . :   0
                                  IPCS IOAs  . . . . . :   0

Select one of the following:

     1. Change CPU and other resource values
     2. Work with disk resources
     3. Edit ASPs
     4. Edit main storage pools
     5. Work with communications resources

Selection or command
===> 1
F3=Exit   F4=Prompt   F9=Retrieve   F12=Cancel   F13=Check configuration
F17=Correct configuration          F24=More keys
```

Figure 110. Configuration

Select the user defined AS/400 partition model, **S40-LP01** (4 processors), based on 9406-S40 #2207 8-way server using option **1**.

```
                                Select CPU

Type option, press Enter.
  1=Select

                    System   Archi-  --Relative Perf---   Number of  Max Stor
Opt  CPU Model      Unit   tecture    Normal    Server  Processors    (MB)
          B10       9404    *CISC      1.00                    1         16
 1   S40   LP01    9406    *RISC     15.00    500.00          4      40960
     20S   2010    9402    *CISC      1.90      5.90          1        128
     200   2030    9402    *CISC      2.50                    1         24
     200   2031    9402    *CISC      4.00                    1         56
     200   2032    9402    *CISC      5.80                    1        128
     300   2040    9406    *CISC      4.00                    1         72
     300   2041    9406    *CISC      5.80                    1         80
     300   2042    9406    *CISC      7.20                    1        160
     310   2043    9406    *CISC     10.90                    1        832
     310   2044    9406    *CISC     19.00                    2        832
     320   2050    9406    *CISC     23.20                    1       1536
     320   2051    9406    *CISC     41.30                    2       1536
                                                                      More...
F12=Cancel
```

Figure 111. Select CPU

The S40-LP01 is created from the AS/400 server, Model 9406-S40 #2207 8-way processor. The S40-LP01 has only four processors with a relative normal performance (interactive) rating of 15 and relative server performance rating of 500.

```
           System   Archi-  --Relative Perf---   Number of  Max Stor
CPU  Model  Unit   tecture    Normal    Server  Processors    (MB)

S40  2207  9406    *RISC     30.41    998.55       8        40960
```

Note

BEST/1 uses its own performance rating figure. It is not the CPW number.

You can project the combined workloads on the user defined model, S40-LP01. The result is shown in Figure 112.

```
                        Display Analysis Summary

                         Stor    CPU  -Disk IOPs--  -Disk Ctls--  -Disk Arms--
Period     CPU Model     (MB)   Util   Nbr    Util   Nbr    Util   Nbr    Util
Base       S40 LP01      2048   73.5     5    13.5     9     2.4    70     5.2
+10%       S40 LP01      2048   80.8     5    14.9     9     2.7    70     5.8
+10%       S40 LP01      2048   88.9     5    16.4     9     2.9    70     6.4
+10%       S40 LP01      2048   97.8     5    18.0     9     3.2    70     7.2
+5%        S40 LP01      2048   99.0     5    18.1     9     3.2    70     7.2

                                                                         Bottom
           ----Inter Rsp Time----  -------Inter--------  -----Non-Inter------
Period      Local    LAN     WAN  CPU Util     Trans/Hr  CPU Util     Trans/Hr
Base           .7    4.3      .0       .0           95     73.4     28064038
+10%           .7    4.3      .0       .0          104     80.8     30870438
+10%           .7    4.3      .0       .0          114     88.9     33957484
+10%           .7    4.4      .0       .0          126     97.8     37353236
+5%            .7    4.3      .0       .0          132     99.0     37619880


                                                                         Bottom
F3=Exit   F10=Re-analyze   F11=Alternative view        F12=Cancel
F15=Configuration menu     F17=Analyze multiple points  F24=More keys
```

Figure 112. Projected Batch Performance for the Combined Workload

To simplify the situation, we made the assumption that the workloads on the two machines are exactly similar and each one of them drives the 9406-53S #2157 4-way processor to 98.4% with 14,032,018 transactions per hour. This allows us to directly combine the batch runtimes of the two workloads on the partition with four processors of a 9406-S40 #2207 8-way server.

The projected runtime of the combined workload is:

```
4 hours x ((14,032,018 + 14,032,018)/ 37,619,880) = 3 hours.
```

The batch CPW of the S40-LP01 is:

```
3660 x (500 / 998.55) = 1833.
Where 3660 is the 9406-S40 #2207 processor's C/S CPW.
```

If the customer accepts the current batch window, four hours, then the required batch performance for the combined workload is 1,300 (650 x2) batch CPW. However, if the customer wants to shorten the batch window from four hours to three hours, then the required batch performance is 1,833 batch CPW.

Appendix F. Special Notices

This publication is intended to help IBM Business Partners and IBM Global Service Professionals to prepare and plan for implementing Logical Partitioning on AS/400. The information in this publication is not intended as the specification of any programming interfaces that are provided by the AS/400 system. See the PUBLICATIONS section of the IBM Programming Announcement for the AS/400 for more information about what publications are considered to be product documentation.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

IBM may have patents or pending patent applications covering subject matter 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 the IBM Director of Licensing, IBM Corporation, 500 Columbus Avenue, Thornwood, NY 10594 USA.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact IBM Corporation, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The information about non-IBM ("vendor") products in this manual has been supplied by the vendor and IBM assumes no responsibility for its accuracy or completeness. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have

been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any pointers in this publication to external Web sites are provided for convenience only and do not in any manner serve as an endorsement of these Web sites.

Any performance data contained in this document was determined in a controlled environment, and therefore, the results that may be obtained in other operating environments may vary significantly. Users of this document should verify the applicable data for their specific environment.

The following document contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples contain 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.

Reference to PTF numbers that have not been released through the normal distribution process does not imply general availability. The purpose of including these reference numbers is to alert IBM customers to specific information relative to the implementation of the PTF when it becomes available to each customer according to the normal IBM PTF distribution process.

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

CT	DB2
DRDA	GDDM
IBM ®	ImagePlus
Netfinity	OfficeVision
OS/2	OS/400
RS/6000	S/370
Service Director	SP
System/38	System/390
VisualAge	VisualInfo
XT	400

The following terms are trademarks of other companies:

C-bus is a trademark of Corollary, Inc. in the United States and/or other countries.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and/or other countries.

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

PC Direct is a trademark of Ziff Communications Company in the United States and/or other countries and is used by IBM Corporation under license.

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States and/or other countries.

UNIX is a registered trademark in the United States and/or other countries licensed exclusively through X/Open Company Limited.

SET and the SET logo are trademarks owned by SET Secure Electronic Transaction LLC.

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

Appendix G. Related Publications

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

G.1 International Technical Support Organization Publications

For information on ordering these ITSO publications see “How to Get ITSO Redbooks” on page 211.

- *AS/400e System Handbook V4R4*, GA19-5486
- *AS/400 System Builder V4R4*, SG24-2155
- *AS/400 Server Capacity Planning*, SG24-2159
- *AS/400 Consolidation Strategies and Implementation*, SG24-5186
- *AS/400 Remote Journal Function for High Availability and Data Replication*, SG24-5189

G.2 Redbooks on CD-ROMs

Redbooks are also available on the following CD-ROMs. Click the CD-ROMs button at `http://www.redbooks.ibm.com/` for information about all the CD-ROMs offered, updates and formats.

CD-ROM Title	Collection Kit Number
System/390 Redbooks Collection	SK2T-2177
Networking and Systems Management Redbooks Collection	SK2T-6022
Transaction Processing and Data Management Redbooks Collection	SK2T-8038
Lotus Redbooks Collection	SK2T-8039
Tivoli Redbooks Collection	SK2T-8044
AS/400 Redbooks Collection	SK2T-2849
Netfinity Hardware and Software Redbooks Collection	SK2T-8046
RS/6000 Redbooks Collection (BkMgr)	SK2T-8040
RS/6000 Redbooks Collection (PDF Format)	SK2T-8043
Application Development Redbooks Collection	SK2T-8037

G.3 Other Publications

These publications are also relevant as further information sources:

- *AS/400 Road Map for Changing to PowerPC Technology V4R3*, SA41-5150
- *System Upgrade Roadmap* (RISC to RISC) V4R3, SA41-5155
- *OS/400 Distributed Data Management*, SC41-3307
- *Client Access for Windows 95/NT Setup V3R2*, SC41-3512
- *Basic System Operation, Administration, and Problem Handling*, SC41-5206
- *Backup and Recovery*, SC41-5304
- *Performance Tools for AS/400*, SC41-5340
- *BEST/1 Capacity Planning Tool V4R2*, SC41-5341
- *OptiConnect for OS/400*, SC41-5414
- *Operations Console Setup*, SC41-5508
- *AS/400 System API Reference V4R3*, SC41-5801

How to Get ITSO Redbooks

This section explains how both customers and IBM employees can find out about ITSO redbooks, redpieces, and CD-ROMs. A form for ordering books and CD-ROMs by fax or e-mail is also provided.

- **Redbooks Web Site** http://www.redbooks.ibm.com/

  Search for, view, download, or order hardcopy/CD-ROM redbooks from the redbooks Web site. Also read redpieces and download additional materials (code samples or diskette/CD-ROM images) from this redbooks site.

  Redpieces are redbooks in progress; not all redbooks become redpieces and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

- **E-mail Orders**

  Send orders by e-mail including information from the redbooks fax order form to:

	e-mail address
In United States	usib6fpl@ibmmail.com
Outside North America	Contact information is in the “How to Order” section at this site: http://www.elink.ibmlink.ibm.com/pbl/pbl/

- **Telephone Orders**

United States (toll free)	1-800-879-2755
Canada (toll free)	1-800-IBM-4YOU
Outside North America	Country coordinator phone number is in the “How to Order” section at this site: http://www.elink.ibmlink.ibm.com/pbl/pbl/

- **Fax Orders**

United States (toll free)	1-800-445-9269
Canada	1-403-267-4455
Outside North America	Fax phone number is in the “How to Order” section at this site: http://www.elink.ibmlink.ibm.com/pbl/pbl/

This information was current at the time of publication, but is continually subject to change. The latest information may be found at the redbooks Web site.

IBM Intranet for Employees

IBM employees may register for information on workshops, residencies, and redbooks by accessing the IBM Intranet Web site at http://w3.itso.ibm.com/ and clicking the ITSO Mailing List button. Look in the Materials repository for workshops, presentations, papers, and Web pages developed and written by the ITSO technical professionals; click the Additional Materials button. Employees may access MyNews at http://w3.ibm.com/ for redbook, residency, and workshop announcements.

IBM Redbook Fax Order Form

Please send me the following:

Title	Order Number	Quantity

First name	Last name	
Company		
Address		
City	Postal code	Country
Telephone number	Telefax number	VAT number

☐ Invoice to customer number _____

☐ Credit card number _____

Credit card expiration date	Card issued to	Signature

We accept American Express, Diners, Eurocard, Master Card, and Visa. Payment by credit card not available in all countries. Signature mandatory for credit card payment.

Index

Symbols

Numerics

A

B

C

D

E

F

H

I

J

L

M

N

O

P

Q

R

S

T

U

V

W

ITSO Redbook Evaluation

Slicing the AS/400 with Logical Partitioning: A How to Guide
SG24-5439-00

Your feedback is very important to help us maintain the quality of ITSO redbooks. **Please complete this questionnaire and return it using one of the following methods:**

- Use the online evaluation form found at `http://www.redbooks.ibm.com/`
- Fax this form to: USA International Access Code + 1 914 432 8264
- Send your comments in an Internet note to `redbook@us.ibm.com`

Which of the following best describes you?
_ Customer _ Business Partner _ Solution Developer _ IBM employee
_ None of the above

Please rate your overall satisfaction with this book using the scale:
(1 = very good, 2 = good, 3 = average, 4 = poor, 5 = very poor)

Overall Satisfaction _____

Please answer the following questions:

Was this redbook published in time for your needs? Yes___ No___

If no, please explain:

What other redbooks would you like to see published?

Comments/Suggestions: (THANK YOU FOR YOUR FEEDBACK!)

SG24-5439-00
Printed in the U.S.A.

Slicing the AS/400 with Logical Partitioning: A How to Guide

SG24-5439-00