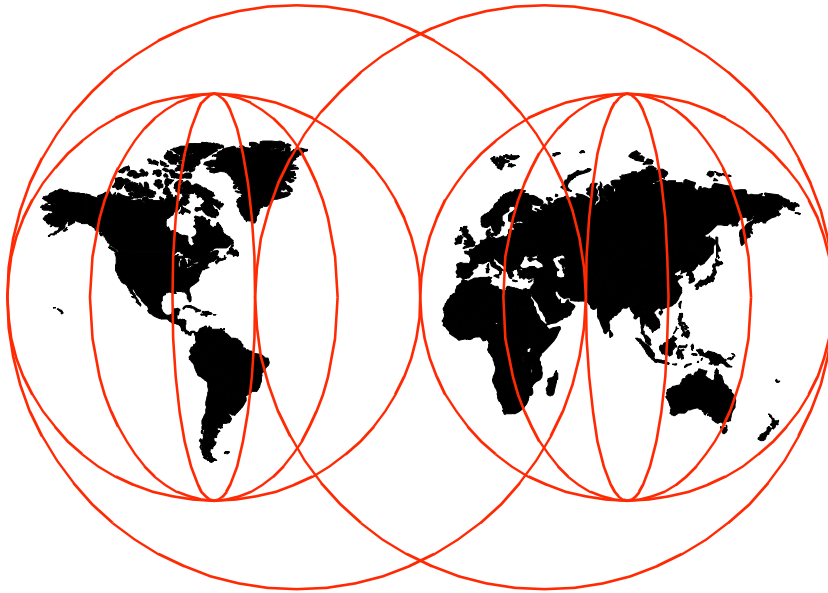Lotus

IBM

High Availability and Scalability with Domino Clustering and Partitioning on AIX

Marcelo R. Barrios, Ole Conradsen, Charles Haramoto, Didac Marin

International Technical Support Organization

http://www.redbooks.ibm.com

SG24-5163-00

IBM

International Technical Support Organization

**High Availability and Scalability with
Domino Clustering and Partitioning on AIX**

August 1998

IBM

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix C, "Special Notices" on page 149.

First Edition (August 1998)

This edition applies to Lotus Domino Version 4.6.1, for use with the AIX Version 4 operating system.

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. HYJ Mail Station P099
522 South Road
Poughkeepsie, New York 12601-5400

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

Lotus Domino servers are an integral part of the IBM Internet technology. The data managed by these servers must be highly available in a world where online information is the corner stone of business success.

This redbook gives a comprehensive explanation of the different ways of making a Domino server highly available. It describes the Domino Clustering technology that is part of the Enterprise Server offering of Lotus Notes servers, as well as the standard HACMP implementation for AIX and the RS/6000 SP.

The redbook will help you to plan, install, and configure a highly available Domino server using Domino Clustering, including Partitioning, and HACMP technology, including the special version developed for the RS/6000 SP.

This book offers a good deal of information about what kind of environment will best fit your needs for high availability and scalability, as well as detailed examples for implementing these different environments. We have also included some tests we ran to calculate the performance impact of using the Domino Clustering technology along with complementary solutions such as Partitioning and HACMP.

Finally, we have included several references to online and paper documentation for further investigation of high availability and scalability with Domino Clustering and Partitioning on AIX.

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 Poughkeepsie Center.

Marcelo R. Barrios is a project leader at the International Technical Support Organization, Poughkeepsie Center. He has been with IBM for five years working in different areas related to RS/6000. Currently he focuses on RS/6000 SP technology by writing redbooks and teaching IBM classes worldwide.

Ole Conradsen is a Lotus Notes and AIX System Architect in IBM Denmark. His area of expertise is integration of Notes and AIX with other systems, especially Web-based solutions. He has worked with AIX since 1988 and Lotus Notes for

three years. He holds a Master of Science degree in Electrical Engineering from Danmarks Tekniske Universitet.

Charles Haramoto is an IT Architect at IBM Global Services in Chile. He has been with IBM for eight years working in systems engineering, networking services and consulting. Now he focuses on Lotus Domino groupware technologies for implementing network computing and e-business solutions.

Didac Marin is an IBM-certified AIX Advanced Technical Expert and Specialist on the RS/6000 SP. He has a degree in Computer Science in Facultat d'Informatica de Barcelona, at the Polytechnic University of Catalunya. During his third and fourth years at the University, he worked at Hewlett-Packard with HP-UX and MPE XL in the PSO department in Barcelona. For the last four years, Didac worked in ISASA (Integracion de Sistemas Abiertos S.A.) with AIX, PSSP and everything related to these environments. Didac is a co-author of the redbook *GPFS: A Parallel File System*, SG24-5145.

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

Ginny Roarabaugh
IBM Austin

Mike Kistler
IBM Austin

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 179 to the fax number shown on the form.

- Use the electronic evaluation form found on the Redbooks Web sites:

  For Internet users   `http://www.redbooks.ibm.com`
  For IBM Intranet users   `http://w3.itso.ibm.com`

- Send us a note at the following address:

    `redbook@us.ibm.com`

Chapter 1. Introduction to Domino Clustering

Clustering has come to mean different things to different people. By definition, *clustering* refers to a loosely-coupled set of systems and/or application instances that are coordinated to provide higher availability and scalability, and enable more efficient use of resources than is possible from a single application instance and/or system.

Domino Clustering refers to a set of Domino servers providing load balancing and failover capabilities through event-driven replication cross platforms. Domino Clustering is part of the new Lotus Domino Enterprise Server available with release 4.6.2 of Domino Server for general usage.

The use of Domino Clustering has the following advantages:

- Failover - With failover, users can continue accessing their critical data even after a server becomes unavailable. When a server goes down, Domino Clustering and the cluster-aware Notes client help redirect user requests to the best available data source in the cluster.

- Workload Balancing - With the server availability threshold parameter, the administrator can redirect users accessing a heavily loaded server to another more available server. This allows for workload balancing and efficient resource utilization in a cluster. You can add servers to your cluster as your workload requirements increase.

- Cluster Replication - Domino clustering provides event-driven cluster replication among servers in the cluster to keep tight synchronization of data across the cluster.

- Scalability - As your needs increase, you can add servers to your cluster and add replicas of critical information throughout your cluster in order to distribute workload.

1.1 Components of a Domino Cluster

As with most things in Domino, a server cluster has its roots in the Public Address Book (NAMES.NSF). The Public Address Book is where you create a cluster, and then add or delete servers from that cluster. When you add a server to a cluster, the Administration Process (ADMINP) adds the Cluster name to the Server document.

For the actual day-to-day operations of the cluster, Domino uses the following components:

1

- Cluster Administration Process (CLADMIN) - This server task is responsible for the correct start-up of all cluster components. On clustered servers, the process runs automatically at server start-up and whenever the cluster membership changes.

- Cluster Analysis - Available from the administration panel, this tool enables you to analyze the configuration of a cluster and determine whether it has been set up correctly. By default, it stores the results in the Cluster Analysis database (CLUSTA4.NSF).

- Cluster Manager - The Cluster Manager tracks all members in a cluster. It keeps a list of which servers in the cluster are currently available and maintains information about the workload on each server.

- Cluster Replicator (CLREPL) - This server task is responsible for synchronizing databases and their replicas in a cluster. The Cluster Replicator uses the Cluster Database Directory to determine which local databases have replicas on the cluster members. Then, whenever a change occurs to a database, the Cluster Replicator pushes the change to the other replicas in the cluster.

1.2 High Availability with Domino Clustering

In networked computing systems, high availability is achieved by providing synchronized redundancy of critical system components by clustering them. As previously noted, a cluster is a collection of hardware or software components that clients on the network access as a single, virtual resource that is highly available. Individual cluster components can be physically located in the same room or dispersed around the world and connected by a network that spans the enterprise. With clusters, key network elements are not only duplicated, but also connected to work together as a team. This enables the synchronized redundant components to automatically and transparently take over for failed components and maintain data availability.

Synchronization of the cluster components is the key to this style of high availability. For example, with Domino clusters, if one database is updated and its replica remains unchanged, the two databases are no longer duplicates. The databases must establish frequent communication over the cluster network to stay in synch. When databases on different servers are identical, they are said to be "synchronized".

However, high availability is not only about hardware and basic software availability; it also has to include service or application availability in order to provide a real solution. Lotus Domino provides this type of high availability

through its Domino Clustering feature, which is included in the Domino Enterprise Server.

Domino clusters are interconnected teams of Domino servers that provide uninterrupted access to Notes networked information resources, including Notes messaging, databases, and other service components.

Using Domino clustering, you can connect up to six Domino servers in the same Notes domain with a local area network. There is no requirement regarding the type of operating system or hardware platform each Domino server runs.

1.3 Workload Balancing with Domino Clusters

Key requirements for customers using Domino for enterprise-class, business-critical applications are scalability and good resource utilization. Basically, scalability is the ability to add computing power to an existing system in a seamless fashion. A key aspect of scalability is *workload balancing*, which is the ability to distribute workload across a clustered system in a way that maximizes the resource utilization.

The Domino server and Notes client work together to provide workload balancing. When running as part of a cluster, the Domino server constantly monitors its own workload. To measure the workload, the Cluster Manager on the server monitors the average response time of a representative set of server operations initiated by Notes clients (network time is not considered). The Cluster Manager also polls all the other servers in the cluster to determine their workload. When the workload on a server exceeds a certain designated level, the server becomes "busy", and the Domino server rejects subsequent database open requests until the workload falls back below the specified level.

When the cluster-aware client (Notes R4 or later) tries to access a database on a busy server, it receives an error code indicating that the server is busy. The client then contacts the Cluster Manager on one of the servers in the cluster. The Cluster Manager uses the Cluster Database Directory (CLDBDIR) to determine which other servers in the cluster have replicas of the database being requested, and then selects the least heavily loaded server to handle the client request. The client then reissues the open request to this new server.

1.3.1 Distribution of Databases in the Cluster

In a cluster, the distribution of users and databases takes on a new importance. When a server in the cluster fails, user requests are automatically redirected to other servers in the cluster. Ideally, this load should be spread equally across all other servers in the cluster. However, this can only happen when replicas of the databases of the failed server are spread roughly equally across the other servers in the cluster.

An example can illustrate this best. Suppose you have 1200 mail users that you want to put on a cluster with four servers. To start, you will probably allocate 300 users to each server. Now, to give these users high availability to their mail databases, you want to create a replica of each user's mail file on another server in the cluster. You might take all users on Server 1 and put a replica of their mail file on Server 2. This is not a good idea. If Server 1 fails, all 300 of its users will be redirected to Server 2. Servers 3 and 4 will not absorb any of this failover load, because the necessary databases are only available on Server 2.

Clearly, a better approach is to spread the replicas for Server 1's users across the other three servers. If these are spread evenly -- that is, 100 of Server 1's users on Server 2, 100 on Server 3, and 100 on Server 4 -- a failure of Server 1 should result in a roughly equal increase in workload for the other three servers in the cluster.

1.3.2 The Server Availability Index

As mentioned in 1.3, "Workload Balancing with Domino Clusters" on page 3, each server in a cluster periodically determines its own workload, based on the average response time of requests recently processed by the server. The workload on the server is expressed as the server availability index, which is a value between 0 and 100, where 100 indicates a lightly loaded server (fast response times), and 0 is a heavily loaded server (slow response times). Despite the fact that the server availability index is a number between 0 and 100, it is not a percentage. Some people think that a server availability index of, say, 85, means that the server is 85% available. This is not the case.

The server availability index is closely related to a common performance metric called the expansion factor. The expansion factor is simply the ratio of the response time for a function under the current load to the response time for this same function in an optimum (light load) condition. So, for example, if the system currently takes 3 seconds to perform a database open, but could perform the same database open in .3 seconds under optimum conditions, the expansion factor for this operation is 10. The expansion factor for a set of operations can be computed as a simple weighted average. To compute the

server availability index, the Domino server computes the expansion factor for a representative set of Notes RPC transactions over a recent time interval (roughly the last minute). The server availability index is then set to 100 minus this expansion factor.

Remember that the server availability index only considers the response time as measured at the server, which is typically only a small portion of the overall response time as seen by clients. In particular, the network time between the client and server often accounts for a significant portion of client response time. So a server availability index of 90 does not indicate that the response time as seen by clients is ten times the optimal value -- only that the server processing of this request took ten times longer than the optimal value.

1.3.3 The Server Availability Threshold

When Domino recalculates the server availability index (approximately once a minute), it checks to see if the index is below the server availability threshold. If the server availability index is less than the server availability threshold, the server is marked as busy. In other words, the server availability threshold specifies the lowest value of the server availability index for which the server should be considered to be available.

You can set the new threshold value by modifying the Notes.ini file, in which case it will be effective the next time the server starts, or you can set it from the server console, in which case the change is effective immediately.

The default value for the server availability threshold is 0, which means load balancing is effectively disabled. Specifying a threshold value of 100 puts the server into the busy state regardless of its actual availability.

1.3.4 Selecting the Proper Server Availability Threshold

As you have probably guessed, the server availability threshold is a key configuration setting for workload balancing. Therefore, you should choose this parameter with some care. Setting the threshold too high can cause user requests to fail unnecessarily. Setting the threshold too low can result in poor performance for some users that may have received better service from another server.

One point which must be stressed is that workload balancing is not a solution for a general capacity problem. If your Domino servers are struggling to keep up with the workload they have, and there are no other available servers to handle the excess workload, enabling workload balancing will only exacerbate the problem. In other words, do not think that increasing the server availability threshold will necessarily make your server more

responsive. If there is nowhere else to send client requests, they will continue to be handled by the busy server, and the process of looking for another available server for each request will only worsen the workload on the server.

To determine the proper value for the server availability threshold, you should start by simply monitoring the server availability index during periods of normal to heavy load. There are a number of ways to do this. One way is to use the built-in statistics monitoring of Domino.

It may seem natural to set the server availability threshold to the same value on all servers in the cluster. While this may be a good rule of thumb, differences in hardware, operating systems, and levels of the Domino server can influence the server availability index and thus the proper setting of the server availability threshold.

Once you have gathered some data on the range of typical values of the server availability index for a server, the next step is to select an initial value for the server availability threshold. This should be a value toward the *lower end* of the range of typical values. You should also consider how a server outage may impact server workload. If a server in the cluster fails, the failover capability in Domino clustering will direct clients to other servers in the cluster. To allow for this case, you may want to set the server availability threshold to allow some "extra" capacity to handle the failover workload. Note that the extra capacity needed for failover depends on how many servers are in the cluster. For a cluster with just two servers, you would need to allow for an almost 100% increase in workload in the event of a server failure. When there are six servers in the cluster, each server would only need to handle a roughly 20% increase in workload.

Once you have selected an initial value, configure this on the server and monitor its operations. Domino gathers a number of statistics on cluster failover and workload balancing that you can use to monitor how well things are going. You can see these statistics by using the Show Statistics server command at the server console. You can also report statistics to any database designed for this purpose, although typically the database is the Statistics database (STATREP.NSF). The Collector or Reporter task creates the Statistics database automatically if you choose to report statistics to it and if it does not exist already. Cluster statistics are available in the Statistics Report/Cluster view.

Chapter 2. Planning for Domino Servers

Planning is the process of preparing for the installation of Domino servers. This involves planning of the standard Domino installation elements domain name, organization name, server names and the naming structure. The planning also includes knowing which type of workload you expect on the server. The requirements for Web servers and mail servers are different. How many users do you expect and what will they be doing? When planning for clustered and partitioned servers, it is important to get these parameters right the first time, because building clusters is based on this information.

Also, before installing the server, you should know if you will use Domino clustering or partitioning, because it must be selected at installation time.

The goal of this chapter is to describe some technologies, what they are intended for, and how to use them to meet customer requirements. These requirements will dictate how to configure the hardware, and how to install and configure Domino.

The following areas need to be considered:

- Applications - You should know the mix of applications on the server to get information on the type of workload, such as: Does a Web server have read-only access? Is it a mail server where the users are sending mail one to one, in which case the write access is maybe just as high as the read access?

- CPU - What is the workload on the system? How many users are using the system? How is the load during the day? What response times are acceptable? What kind of CPU does the customer want, that is, more smaller CPUs or one big one?

- Disk and I/O bandwidth - Based on your knowledge of the application, you should know how much disk capacity is needed, and what the I/O bandwidth requirements are. Remember that Domino puts a heavy load on the I/O system, so you need as many disk drives as possible. You would also need more disk controllers to get a higher bandwidth.

- LAN bandwidth - How are your users connected to the servers? Do you have sufficient capacity? If you are planning to run Domino Clustering, do you have LAN bandwidth and low latency LAN between servers? If you are using Domino Clustering, the SP Switch is a very attractive connection.

- Availability - What are the requirements for availability? Domino clustering is one way to get more availability, but other techniques should be used in

combination with Domino clustering. RAID disks and disk mirroring can give you advantages, such as:

- Data protection and availability - disks can be concurrently repaired while the Domino servers stay up.
- Performance advantage for read-only access to data and file system logs.
- Backup flexibility - with mirrored disks. You can make a mirrored copy of data offline for backup, as opposed to taking a server offline.

If you choose not to mirror all Domino external disks, it is recommended that you mirror at least the Domino server data (notes.ini, name and address books, and file system logs). This requires only one extra disk per server and ensures that even if a disk is lost the server remains available.

Planning an internal disk for Domino executable code

The file system where Domino is installed needs to be 200MB to hold the Domino server code. This file system holds the binaries. The code is in disk cache, so there are no special I/O requirements for this disk. We recommend a separate file system, /usr/lpp/lotus (or expand /usr by 200MB). You can put the binaries on internal disks in rootvg volumengroup, or in a volumengroup made for Domino.

Planning an external disk for the Domino server data directory

For all servers, the Domino Server data directory, where the key databases reside, needs to be large -- at least 2GB. We recommend that you break out the data directory as a separate file system in a separate volume group. This makes it easy to migrate the data disk to another CPU. Based on the server type we can say that for:

- Mail servers

  Allocate a base amount of space for the data directory (notes.ini, names.nsf, and name and address books). Add to that the amount of mail space allocated per user (typically 10-25MB) times the number of users. If the server is used for other databases besides the mail databases, we suggest separating the base data directory and user mail databases, for ease of administration. If you have more than 300 mail users on a mail server, it might be a good idea to split the users into different directories. If you open a folder with 1000 mail databases, the user would wait for several seconds while Domino lists all the directories. However, it requires some extra administration work to separate the users into different subdirectories.

- Application servers

  Allocate a base amount of space for the data directory and add to that the size of the application databases.

- Mail Routing Servers

  The external disk space will be used for mail in flight, which is typically in the tens of gigabytes. Although this space will not be fully utilized, under normal circumstances, remember to allow sufficient space to hold backed-up mail in the event a network connection is lost to a routing destination (or a routing destination failure occurs).

- Replication servers

  Allow sufficient space to hold all of the database files that will be replicated outside the domain.

Because of the way Domino works, the JFS log logical volume(s) are very heavily used, so for performance reasons consider putting these on a disk by themselves.

Disk mirroring or RAID will increase your disk requirements, so be sure to plan accordingly.

2.1 Technologies

This section describes the different technologies available for implementing Domino servers in an AIX or AIX/PSSP environment:

- Domino clustering
- Domino partitioned servers
- HACMP
- HACMP/ES

Domino clustering and Domino partitioned servers are features of the Domino Enterprise Server license, and are available for all AIX platforms (RS/6000 and SP nodes) and AIX/PSSP environments (SP nodes).

HACMP is a product available for all AIX platforms (RS/6000 and SP nodes), while HACMP ES is a product available only for AIX/PSSP environments (SP nodes only).

2.1.1 Domino Clustering

The transparent failover feature or the load balancing feature of Domino Clustering can redirect the open request for a database from an unavailable or overloaded server to another available or unloaded server of the cluster.

In a clustered Domino environment, replication is event-driven, whereas in a standard Domino environment, replication is schedule-driven. For this reason, you must take special considerations into account when planning a clustered environment. Network bandwidth and the total number of replicas per database within the cluster are critical factors that impact the overall performance of the cluster:

- Clustered servers should be interconnected through a high data transfer bandwidth network to allow for tight synchronization between the databases in the cluster.

- Frequently accessed read-mostly databases, such as discussion databases, can be spread across several servers in the cluster; but single owned read-write databases, such as mail databases, can have only one replica for availability purposes.

- Well-designed database documents can reduce the replication traffic between clustered Domino servers. Keep in mind that whenever a field is modified, it has to be pushed to the replicas maintained by the cluster.

Clustering failover and load balancing for non-Notes clients is not supported in the current version of Domino. A workaround to this limitation is IBM's Interactive Network Dispatcher, which allows load balancing for HTTP, SSL, NNTP, POP3 and SMTP traffic; and also allows various load metrics when combined with the Interactive Session Support (ISS), the optional agent code for monitoring controlled systems.

For more information on Domino clustering, refer to the Domino online documentation *Domino Administration Help, Extending the Domino System, Chapter 9 Setting Up a Cluster*; or *Lotus Domino Server Release 4.5 on AIX Systems: Installation, Customization and Administration*, SG24-4694.

2.1.2 Domino Partitioned Servers

Domino partitioned servers are a feature of the Domino Enterprise Server license that allows up to six Domino servers to run on a single computer. Partitioning a single computer into separate Domino servers provides the following:

- A reduced number of computers to own and administer in order to support an independent group of users

- Full Domino security between partitioned servers when implementing independent groups of users

Partitioned servers share the same Domino binaries, but each has its own assigned Domino data directory containing the required set of system files such as Public Address Book, notes.ini and database files associated with each Domino server. Also, each partitioned server has memory requirements that should be sized according to the nature of the tasks that are going to be running in each of them.

Because note every partitioned server requires its own CPU, it is possible to run partitioned servers both in single CPU systems or in SMP systems with multiple CPUs. High-end configuration systems (with powerful single or multiple CPUs and large amounts of memory) are good candidates for implementing partitioned servers, due to a limitation on the ability of a single Domino server to take full advantage of such configurations. This limitation will be resolved in future versions of the Domino server. For performance improvement results, refer to Chapter 4.2, "Performance In Partitioned Domino Servers" on page 101.

Partitioned servers can share a single network interface adapter (by either using a different IP address for the same adapter or using different IP ports for the same IP address), or use multiple network interface cards (by using a different IP address for each adapter).

A partitioned server can also be a member of a Domino cluster. However, for availability reasons, we recommend that you locate the other servers of the cluster in different systems.

For more information on Domino partitioned servers, refer to the Domino online documentation *Domino Administration Help, Extending the Domino System, Chapter 10 Setting Up a Partitioned Server*; or *Lotus Domino Server Release 4.5 on AIX Systems: Installation, Customization and Administration*, SG24-4694.

2.1.3 HACMP

High Availability Cluster Multi Processing (HACMP) for AIX Version 4.2.2 is a control application that can link up to eight RS/6000 servers or SP nodes into a highly available cluster.

HACMP clusters can be configured in several modes for different types of processing requirements:

- Concurrent access mode suits environments where all of the processors must work on the same work load and share the same data at the same time.

- In mutual takeover mode, the processors share the work load and back each other up.

- Idle standby allows one node to back up any of the other nodes in the cluster.

For a Domino server environment, only mutual takeover and idle standby configurations are possible since concurrent access to Domino databases is not supported. For example, in a two-machine scenario with shared disks, one machine can take over the data file systems of the other in case of a failure.

By default, HACMP monitoring and takeover scripts only check for machine or network failures. To incorporate the crash of a Domino server in the scripts already provided with HACMP, customized heartbeat programs and takeover scripts must be included.

For more information on HACMP, refer to *An HACMP Cookbook*, SG24-4553.

2.1.4 HACMP/ES

High Availability Cluster Multi-Processing/Enhanced Scalability (HACMP/ES) for AIX Version 4.2.2 extends the clustering capabilities of HACMP for an SP-specific environment. Up to 16 SP nodes can be linked in the current release.

The functionality provided by HACMP/ES is similar to HACMP. However, the current release of HACMP/ES has important differences that must be considered:

- HACMP/ES uses PSSP Topology Services for heartbeating, PSSP Group Services for membership information, and Group Services and Event Manager for event detection.

  HACMP has its own internal function to detect events and its own heartbeat protocol.

- HACMP/ES is partition-bounded because PSSP Topology Services and Group Services are partition sensitive; for this reason, an HACMP/ES cluster can only run on a single SP system, and cannot be used with High Availability Geographic Cluster (HAGEO) to provide a full solution for disaster recovery.

  HACMP is not partition-bounded, so it can be used with HAGEO to provide disaster recovery for SP systems.

- HACMP/ES cannot dynamically reconfigure the topology-related definitions (nodes, networks and adapters).

  HACMP can change the configuration of a running cluster.

For more information on HACMP/ES, refer to *HACMP Enhanced Scalability*, SG24-2081.

2.2 High Availability

When planning for a highly available Domino installation, it is important to check other elements than the Domino installation itself. You should first consider which part or component of the total system is most likely to fail, and what can be done to avoid the failure.

Most often a server fails or stops service due to manual failures, or the system administrator stops the server without warning and installs new software on the system without proper testing and without realizing the server is running a mission-critical application. New applications are installed without proper testing, and so on. The following are ways to minimize such problems:

- Limit the number of users with physical access to the servers.
- Keep server passwords safe.
- Never upgrade system software on a production server without testing the upgrade first.
- Never upgrade Domino on a production server without testing the upgrade first.
- Do not install applications to a server without proper testing.
- Only authorized administrators should change server settings.

In addition to these basic steps, you should consider which part of the system is now the most likely to fail:

- The network, including routers and WAN lines
- The power supply
- The CPU
- Disk

To prevent service from stopping because of system failure, there are technologies such as HACMP, Domino clustering and, for Internet solutions,

Interactive Network Dispatcher, which provide methods for continue service even in case of hardware and/or software failures.

Although it seems that Domino clustering provides nearly identical function to HACMP, it is not necessary true. There are cases where both HACMP and Domino clustering can be employed in which both provide useful aspects of high availability and/or scalability. For example, Domino clustering does not provide high availability for server addin tasks (such as, fax gateways, pager gateways, or SMTP MTA). HACMP could be used for this aspect of availability, and Domino clustering could be used for failover and workload balancing of client access.

For cases where you have to make a choice between Domino clustering and HACMP, Table 1 shows the preferred solution for different server functions.

Table 1. Selection of Technology for Availability

Domino Server Function	Preferred Availability Solution	Rationale
Mail	Either HACMP or Domino Clustering	Both Domino Clustering and HACMP can be configured to take over in case of failure, so the decision of which solution should be implemented will be based on other factors, such as previous experience or integration with other subsystems you may have installed.
Mail Router	HACMP	Domino Clustering does not failover mail routing functions. HACMP can failover/restart and recover this functionality.
Application	Domino Clustering	Domino Clustering will load balance and failover users to a backup for properly replicated databases.
Out-of-Domain Replication	Domino Clustering or HACMP	Domino Clustering supports failover out-of-domain replication, so the best choice should be Domino Clustering.

Domino Server Function	Preferred Availability Solution	Rationale
Internet	HACMP or Domino Clustering combined with Interactive Network Dispatcher	Domino HTTP/IP addresses are not failed over by Domino Clustering as with HACMP. However, Interactive Network Dispatcher can balance the load among HTTP servers and provide IP address "takeover". Combined with standard Domino, this solution provides highly available HTTP solutions.
Internotes	Domino Clustering	Internotes Servers are part of the advanced services functions/servers that can be failed over.

2.2.1 Domino Clustering for High Availability

You can improve the availability of your applications and data by adding replica copies of your databases and using Domino clustering to provide the *Takeover* function, which lets you receive database service even if the server you are requesting is not available at the moment. Domino handles this transparently to the user. Domino clustering can create up to six copies of any Domino database on up to six servers, but the more copies you make, the more disk space, CPUs and network bandwidth are required. For most implementations you would use two or three copies, not more.

When you are planning the cluster, you must design it to handle the problems you want protection against. A crucial part of the planning is to decide where database copies should be placed. This is because when another copy is added, it needs also to be replicated, and this consumes CPU power and LAN network bandwidth. Therefore, adding another copy of a database should only be done if necessary. When adding a database replica, you should place it on a server on the same LAN as the original server, or, better, provide a local LAN for cluster traffic, not across routers, WANs or other low-speed LAN connections. When you enable cluster replication between servers, a lot of cluster traffic will occur, and if the traffic has to go across routers and WANs, the replication will be too slow and the load on the LAN too high.

When adding another server to the cluster, you should also be aware that cluster replication by default is enabled for all database replicas. Most databases do not require fast cluster replication. In this case, you should disable cluster replication and let the scheduled replication take care of the replication.

Figure 1 on page 16 shows four servers. Three of the servers are members of a cluster, and one server is outside the cluster. The cluster members are server S1, S2 and S3, and the databases are DB1 to DB4. As you can see DB1 has two copies, one in server S1 and one in server S2. This protects DB1 against system failure in system S1 or S2. However, systems S1 and S2 are placed on the same network, and the network therefore is a single point of failure. If LAN1 fails, the data will no longer be accessible to clients. To protect against failure on LAN1, S3 is placed on another LAN segment, and therefore DB2, which has copies on S1 and S3, is also protected against failure on LAN1. If possible, you should provide a local LAN between S1, S2 and S3 for cluster traffic. If it is impossible to provide a LAN for cluster traffic, you may consider running without one. This is normally not recommended, as Notes clustering requires a high bandwidth LAN connection to handle cluster replication traffic. If you have the required bandwidth from LAN1 to LAN2, or if DB2 is a read-only or read-mostly database, you may use a configuration as shown in Figure 1. This will result in higher availability as the LAN is no longer single point of failure. To get high availability is a trade-off between LAN bandwidth, CPU cycles and availability. In the example, one might want to add another copy of DB2 at server S4, but this would raise the cluster replication traffic from server S1 across the network, which would impact overall system performance.
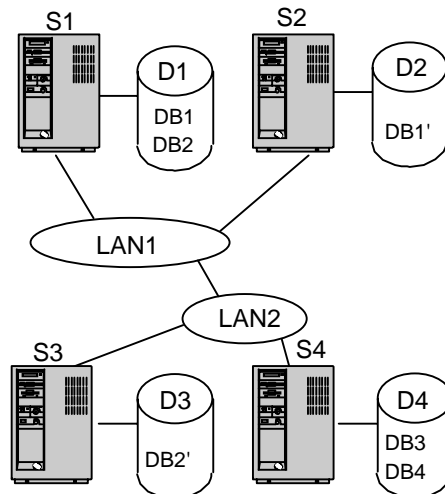
Figure 1. Domino Cluster and ExternalServer.

Domino clustering for high availability has the following advantages:

• Fast and uncomplicated installation.

- Flexibility, because a number of servers can maintain different parts of the cluster.

- It can be used over long distances if LAN bandwidth is available.

2.2.2 AIX HACMP and SP HACMP ES for High Availability

AIX HACMP is the AIX product to enhance system availability. HACMP can control up to eight AIX systems, and prevent failures due to hardware or software problems by system takeover. Figure 2 describes a configuration with two AIX systems, S1 and S2, controlled by HACMP. S1 and S2 share two disks, D1 and D2, where D1 and D2 are mirrors. In normal mode, system S1 has control over D1 and D2, but system S2 also has connection to the disks. If S1 fails, S2 takes over D1 and D2 and continues service.

This configuration means that AIX maintains the copy rather than Domino, the disk copy requires no LAN traffic, and the CPU load for maintaining the copy is minimal compared to the load of the Domino cluster. In this configuration, data is always 100% synchronized, as there is only one database. In a normal production system, S1 owns the disk and Domino runs on S1. In case of a disk failure on disk D1, the system continues operation on disk D2. In case of a system failure on system S1, system S2 activates disks D1 and D2, and the operation could continue, even while system S1 is stopped.
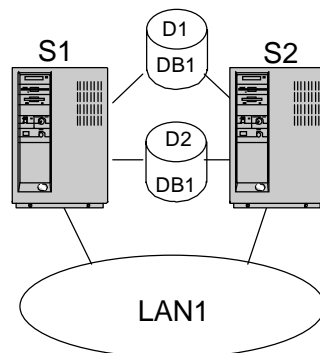
Figure 2. HACMP Cluster

AIX HACMP advantages:

- No extra LAN traffic

- No extra CPU load

- Data is updated in real time, always 100% synchronized for the user

AIX SP HACMP ES is the HACMP product for AIX SP systems. It has advantages over standard HACMP but requires a AIX SP system where HACMP can run on standard IBM RS/6000 hardware. Besides the AIX SP HACMP ES capabilities, the SP system also has a high bandwidth, low latency LAN, and the SP Switch. The switch is also perfect for a Domino cluster if you use the cluster to gain scalability.

2.2.3 Using Interactive Network Dispatcher for High Availability

Interactive Network Dispatcher is a scalable, highly available load-balancing software solution for HTTP, FTP, or other TCP-based servers. It runs on IBM AIX, Windows NT, and Sun Solaris platforms.

With Interactive Network Dispatcher you can load-balance requests to a number of Domino servers even if these are running cross-platform operating systems.

If the Web content is dynamic, Domino clustering combined with Interactive Network Dispatcher can ensure synchronized data and high availability.

Table 2. Interactive Network Dispatcher Features and Benefits

Function	Benefit
High availability	Primary and standby machines are designated, each running Interactive Network Dispatcher. When the primary machine starts up, it and the standby machine immediately synchronize their configuration data. The standby machine remains ready to take over if necessary. Manual switch-over is also possible. With this backup capability (a built-in function of Interactive Network Dispatcher), your clients can always get to the information they need.
Multi-site load balancing	Previously, for balancing traffic across multiple mirror sites, only the servers loads were used as a balancing criteria. Now, with Version 1.2 of Interactive Network Dispatcher, you have the option of using server-to-client proximity (called "ping triangulation") as a balancing criteria.

2.3 Scalability

When you plan for scalability, it is important to distribute the databases, and to make a plan for the databases and their use, you should have an overview

of how much read and write access there is in each database. This is the basic information you need when you decide that a database is a target for clustering. Table 3 shows some database types and typical access to these databases.

Table 3. Typical Read-Write Ratios for Database Types

Database	Read-write Ratios	Typical Access	Description
Mail databases	66% Read 33% Write if each note is read 2 times.	1 user + mailrouter	For scalability, the databases should have one copy only. Scale by adding servers and distributing databases.
Discussion data bases	80% Read 20% write	many	For scalability, this type of database should have replicas on more servers, and a high degree of replication is needed
HTTP servers	95%read 5%write	many	For scalability, this type of database should have replicas on more servers, and a low degree of replication is needed because data changes only occasionally.

When planning the distribution of databases, you should know the use of each database and plan the system to have more copies of read-mostly databases, and one or a few copies of databases with write access.

The mail type of database is accessed only by one user for reading and sending new mail, and the mail router for delivery (write) of new mail. This kind of database with one or few readers/writers should not have replicas. Because of scalability, it might be a good idea to make a replica at some other system to get more availability. This type of database is easy to handle from a Scalability point of view, because it can be migrated to another system. Find the average use of each database and calculate how many databases of this kind each system can support. If this type of database should be replicated for availability reasons, we recommend standard replication, not cluster replication. Systems supporting this type of database should not participate in Domino clusters.

The discussion forums type of database receives input from many users and the documents are read by many users. Each document is read more than once by different users, maybe five times more than written, depending on

the discussion. When somebody puts a new append, everybody should see it directly. Therefore, the database should be replicated or cluster-replicated.

Information databases, where few people update and many read information, are very good for scaling. These databases are not changed often, and are therefore mostly read. Normal information need not be synchronized immediately, and standard replication is sufficient to maintain replicas of databases of this type. Using cluster replication is only recommended for this type of application if data has to be synchronized, because it is not needed and consumes more CPU power than regular replication.

HTTP databases are like information databases, only with more read access. To scale HTTP databases, you should add more replicas and use a product such as Interactive Network Dispatcher to load-balance requests over more servers.

2.3.1 Domino Partitioning and Scaling

Domino partitioning was originally intended for service providers and facility management providers as a way to support more customer environments in one hardware configuration, and it is perfectly suited for this purpose. However, when running Domino on high capacity servers such as IBM RS/6000 and IBM SP servers, it is also possible to get more computing power by using partitioned servers. This is because of changes in configuration of high-end servers. Now these servers are often configured with multiple CPUs (usually SMP architectures) and more than 256MB of real memory; 1GB and 2GB are not uncommon.

Because of the new configurations, it is often possible to scale the Domino environment by use of partitioning. The logic and the rules for distributing databases are the same as for standard Domino servers with respect to CPU, and when increasing the number of servers, you also increase the required I/O bandwidth and need more disks and disk controllers. Be careful when you check a server out and decide if it is a candidate for partitioned server. We have seen, in particular with multi-CPU SMP systems, that even when the CPU is 100% loaded and there is a good balance between user and system CPU usage, you can get substantial improvements by server partitioning.

Figure 3 on page 21 shows a server with four partitions installed, and four separate disk subsystems for Domino databases. You will also need a LAN adapter for each server partition. After installing the partitioned servers, you must define one adapter for each server partition in the notes.ini file.
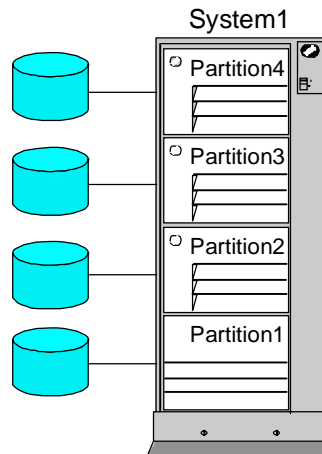
Figure 3. Partitioned Server

When planning for partitioned servers, you should allocate databases so that each server is loaded equally, if possible. You should also look into your disk configuration and see if the capacity is adequate. If database load varies during the day, you should have a mix of databases to maintain equal load.

2.3.2 Scaling with Domino Clustering

As mentioned before, the key to clustering as a method for scaling is to know the use of each database and application. First you should look for read-mostly databases used by many people. These databases would be the number one choice for clustering. Then you should look for databases also used by many people, but for reading and writing. Databases used by one or few users, such as Mail databases, will not necessary gain in performance from clustering, although they will gain in high availability.

These databases should be distributed among servers with only one replica of each database.

Table 4. Database Types and Their Suitability for Clustering

Database Type	Access Type	Number of Users	Scaling with Domino Clustering
Mail	Read and write	One user	No
Discussion	Read-mostly	Many	Yes
Information/ HTTP	Read	Many	Yes

2.4 Performance

This section describes how to plan for performance when using Domino partitioned servers or Domino clustering, when to use each technique, and what factors to consider in your plan.

2.4.1 Domino Partitioned Servers

Domino partitioned servers can increase the performance of a computer when the computer has a high-end configuration and a single Domino server is not able to utilize all its capacity.

As explained in Chapter 2.1, "Technologies" on page 9, this is due to current Domino server limitations. More than one instance of the Domino server can now take full advantage of computers configured with one or more CPUs and large amounts of memory.

2.4.1.1 Domino Considerations

When running partitioned servers, you have to optimize the following Domino parameters:

- Restrict the number of active users.

  Use the *Server_MaxUsers* setting to limit the number of active users on a server.

- Restrict the number of concurrent transactions.

  Use the *Server_Max_Concurrent_Trans* setting to limit the number of concurrently scheduled transactions on a server.

  Lotus recommends that the sum of the limits for concurrent transactions on all partitioned servers must be 20 transactions or less. For example, if two partitioned servers are running, then set the maximum number of concurrent transactions on each server to 10.

2.4.1.2 Hardware Considerations

When planning Domino partitioned servers in high-end computers, consider the following guidelines:

- Though each partitioned server does not require its own CPU, having multiple CPUs is an advantage for running multiple instances of Domino servers.

- Each additional partitioned server requires a minimum of 64MB of memory.

- Each partitioned server requires disk space for its own Domino data directory. All the partitioned servers share the same binary files.

2.4.1.3 Network Considerations

To define the TCP/IP configuration for the partitioned servers, you can use the following:

- Unique IP addresses

  It is possible to use a single NIC or multiple NICs to implement unique IP addresses for each partitioned server.

  As you detect bottlenecks in the NIC or LAN, add more NICs or LAN segments to your machine.

- Unique IP port numbers

  It is possible to implement unique IP port numbers using one NIC for each partitioned server. Assign and configure which of the partitioned servers is going to be the port mapper for the rest, and remove all user workload from this partitioned server. Keep in mind that all the user requests are going to come to this partitioned server before they are routed to the rest of them.

For more information on Domino partitioned servers, refer to *Domino Administration Help*, or *Lotus Domino Server Release 4.5 on AIX Systems: Installation, Customization and Administration*, SG24-4694.

2.4.2 Domino Clustering Considerations

Domino clustering can increase the overall performance of a Domino site when a computer is already at its maximum capacity and has no available resources to allocate in order to increase its throughput, or the number of users that can connect to it.

Domino clustering allows two or more computers to share the user workload. You can spread your current users to more than one computer in order to increase the overall throughput; or you can increase the number of users as you add more computers to the cluster, maintaining the current throughput.

2.4.2.1 Domino Considerations

Use different replication policies depending on the use of each Domino database, as follows:

- Frequently accessed read-mostly databases

  You can have a copy of this kind of database in each of the servers in the cluster since the replication traffic they generate should be low.

- Single owned read-write databases

  Typically these are Mail databases. For availability purposes, at least one replica should be located in a different server, so that replication traffic will be generated as a user updates this kind of database. Avoid having more replicas, since traffic increases as each replica has to be pushed from the modified database to the rest of the replicas.

2.4.2.2 Network Considerations

A high-bandwidth network is recommended for the cluster replication traffic. Plan the replication policy for cluster members according to the bandwidth you have available in your network, as follows:

- LAN Topology

  When the cluster members are physically close, consider an isolated high-bandwidth network for the cluster replication traffic.

  A good example of high bandwidth is the integrated high-speed switch available with SP systems.

- WAN Topology

  When some of the cluster members are connected through a WAN, it is recommended to use T1 links (1.5Mb/s) or better.

  Replace the cluster replication with scheduled replication during periods of lower workload, for all non-critical databases.

For more information on Domino clustering, refer to *Domino Administration Help*; or *Lotus Domino Server Release 4.5 on AIX Systems: Installation, Customization and Administration*, SG24-4694.

Chapter 3. Implementation of Domino Servers

The purpose of this chapter is to discuss the steps for installing and implementing Domino 4.6.1 in the following environments:

- RS/6000 running AIX 4.3
- SP Systems running AIX 4.3 and PSSP 2.4

The items covered are the following:

- Installation steps of Domino servers
- Implementation steps of Domino servers using one of the following technologies:
  - Domino partitioning
  - Domino clustering
  - HACMP for AIX
  - HACMP/ES for SP systems

Since the installation and implementation steps for SP systems are almost the same as for AIX, only the platform-specific differences are mentioned.

3.1 Installation of Domino 4.6.1 on AIX

Domino 4.6.1 installation and configuration on AIX and on SP systems is covered in this topic. It also lists the system requirements for the AIX version of Domino 4.6.1.

The installation process is designed for a future clustered or partitioned Domino configuration. Some of the steps shown are not strictly necessary, but provide a clean and stable initial installation ready to grow to a clustered or partitioned configuration.

For more information on general Domino server installation, refer to the redbook *Lotus Domino Server Release 4.5 on AIX Systems: Installation, Customization and Administration*, SG24-4694.

The following information specifically relates to the AIX operating system.

3.1.1 Hardware and Software Requirements

To run a Domino 4.6.1 Server in an AIX system, the following is required:

- An IBM RS/6000 running AIX 4.1.5, 4.2.1, 4.3.0 or greater.

- A minimum of 64MB of RAM memory, 128MB or more recommended.
- Set up the paging space according to the amount of real memory of your system. Initially follow the general recommendations for AIX. These recommendations are:
  - For systems with a RAM between 64MB and 256MB, the paging space should be: RAM size * 2.
  - For systems with more than 256MB of RAM, use the following equation:

    Paging Space = 512 + (RAM - 256) * 1.25

  You will need to change your paging space according to the real workload your Domino server will handle. We recommend that you monitor the paging space usage in your system to set it up as correctly as possible. To monitor the paging space, use the command `lsps -a`.
- 180MB of disk space for the installation directory.
- 100MB minimum disk space for data. The amount of disk space needed for the data directory depends on your requirements. A base to calculate the disk space needed for Domino mail servers could be 30-50MB per mail user.

Note: Disk space for Domino binaries and data may differ depending on the features you select to install. See Table 5 on page 26 for details.

Table 5. Disk Space Requirements for Domino 4.6.1

Data Directories	Disk Space (Data F.S.)
User Data Files	1MB
Domino Partitioned Server	1MB
Advanced Services Data Files	1MB
Domino Server Data Files	24MB
Notes Workstation Data Files	17MB
Help Lite Database	11MB
Help Database	46MB
Documentation Databases	12MB
TOTAL	**113MB**

- Protocols supported: TCP/IP, X.PC and IPX/SPX.

SPX support for AIX 4.2 is provided by the fileset `ipx.rte`, which comes with the AIX V4.2 Server CD.

SPX support for AIX 4.3 is provided by the fileset `ipx.base`, which comes with the AIX V4.3 Server CD.

Domino Advanced Services clusters and partitioned server configurations do not support the IPX/SPX protocol in Release 4.6.1. At this time, Lotus does not plan to provide IPX/SPX network support for future releases of these features.

- A graphic display supported by the version of AIX you are running (or PC X-Windows emulator).
- A mouse.

3.1.2 Operating System Level

At the time we wrote this redbook, no patches were needed for AIX 4.3.The following AIX 4.1.5 and 4.2.1 filesets must be updated to the specified levels to run Domino 4.6.1.

Table 6. Table of AIX 4.1.5 Patches for Domino 4.6.1

Patch Number	Fileset
U446706	X11.Dt.lib 4.1.5.2
U446812	bos.net.nfs.client 4.1.5.1
U446810	bos.rte.lvm 4.1.5.3
U446245	bos.adt.prof 4.1.5.1
U446855	bos.rte.man 4.1.5.2
U446244	bos.rte.libc 4.1.5.1
U446761	bos.rte.tty 4.1.5.2
U446701	devices.pci.00100100.rte 4.1.5.1
U445109	devices.isa_sio.sa.rte 4.1.5.1
U445116	devices.pci.00100f00.rte 4.1.5.1
U445081	devices.pci.14101800.rte 4.1.5.1
U446247	devices.pci.14105400.X11 4.1.5.1
U446708	devices.pci.14105400.rte 4.1.5.2
U445070	devices.pci.14105e00.X11 4.1.5.1

Patch Number	Fileset
U446745	devices.pci.14105e00.rte 4.1.5.2
U445032	devices.pcmcia.serial.com 4.1.5.1
U449238	xlC.rte For Notes 4.1x and 4.5x use version 3.1.3.0 or 3.1.4.0 For Notes 4.6 use version 3.1.4.4
U446817	If installed, the following fileset needs to be updated to indicated level: bos.rte.mp - 4.1.5.3
U446869	If installed, the following fileset needs to be updated to indicated level: bos.rte.up - 4.1.5.3
U442271	ipx.rte 4.1.5.0
U447411	connect.server.com 4.1.5.3
U447415	connect.protocols 4.1.5.1
U447397	connect.server.nwserve 4.1.5.2
APAR IX72217	Required if using Yellow Pages for name resolution. bos.rte.libc

Table 7. Table of AIX 4.2.1 Patches for Domino 4.6.1

Patch Number	Fileset
U449238	xlC.rte For Notes 4.1x and 4.5x use version 3.1.3.0 or 3.1.4.0 For Notes 4.6 use version 3.1.4.4
U448806	If installed, the following fileset needs to be updated to the indicated level: bos.mp - 4.2.1.0

Patch Number	Fileset
U448841	If installed, the following fileset needs to be updated to indicated level: bos.up - 4.2.1.0
APAR IX72217	Required if using Yellow Pages for name resolution: bos.rte.libc.

To get the latest information of the AIX patch requirements for Domino 4.6.1go to:

```
http://orionweb.lotus.com
```

and search for Patch Requirements for Domino 4.6.

In this Web site you can also find a lot of useful information about Domino setups, configurations, and known problems with their solutions.

3.1.3 Operating System Parameters

Some operating system parameters should be changed for Domino 4.6.1:

- Set the Maximum number of processes allowed per user to 1024. Use the following command:

  ```
  chdev -l sys0 -a maxuproc='1024'
  ```

  This OS parameter must be set to at least 100, but we recommend that you set it to 1024 to cover all possible Domino server installations.

- To enable support for > 1GB databases on AIX, edit the /etc/security/limits file and change the value for fsize to -1 (which means no limit).

- If you will have an entry to start your Domino server in the file /etc/inittab, set the system to automatically reboot after a crash. It makes the restart time of the Domino server shorter. Issue the command:

  ```
  chdev -l sys0 -a autorestart=true
  ```

- To enable support for Notes databases greater than 1GB, set the fsize parameter for user notes in the file /etc/security/limits to `-1` (which means unlimited).

- Set up the size of the /var file system to allow it to hold a system dump. It is very important to capture the dump image when a problem happens. Use the command `sysdumpdev -e` to approximate the free space needed in

the /var file system. The result of the `sysdumpdev` command can vary depending on the use of the system memory when you run the command. The best time to run it is when your Domino server is holding a heavy load.

- Install the fileset `bos.adt.debug`, which contains the dbx debugger. dbx is used by the *Notes System Dump* (nsd) script supplied by Lotus for providing debug information when reporting problems to Lotus support.

3.1.4 Preparing the Environment

Use the following guidelines to prepare your AIX environment. They will give you better performance and flexibility when running your Domino server.

3.1.4.1 Disk Technology

To get the best possible performance and flexibility, and easy management, we recommend that you use SSA disks as external storage for the Domino databases.

If you are using disk drives with capacities greater than 2GB, set the partition size to a value greater than 4MB to use its full capacity.

3.1.4.2 Creation of the Domino File Systems

For the creation of the Domino file system, there are some performance and system management considerations you should take into account.

Performance considerations

- To achieve better performance, you should create the file system for the Domino databases on top of the striped logical volume (if your disk configuration allows you to do this).

- We recommend that you place the Domino binaries and the Domino databases in separate volume groups, allowing parallel access to binaries and data and keeping two different journal file system logs. In case you have binaries and databases in the same volume group, we recommend that you place them in separate file systems and in different disks.

- If you have a separate volume group for the Domino databases with more than one disk, you should place the Log logical volume (which holds the journal file system log) used for the Domino databases file systems as a separate physical disk.

Easy management and flexibility considerations

- It is possible to place the Domino binaries in the AIX file system /usr (without creating a new file system), but this option will complicate your system management.

- Creating the Domino binaries and the Domino databases in separate volume groups will make their management easier. You will be able to apply different volume group maintenance policies (backup, management and so forth). If you cannot place them in separate volume groups, at least try to keep them in separate file systems and different disks.

- Placing your Domino file systems on external disks make easier it to move them from one system to another.

3.1.4.3 About Network Configuration

The minimum requirements to connect a Domino server to the network are:

- Configure at least one network interface in your system to allow client connections.

- Resolve the name of the Domino server.

We give you two examples of how to resolve the name of the Domino server for TCP/IP:

- Using local name resolution, add the following line to the /etc/hosts file:

  ```
  192.168.4.5 server01
  ```

  If you use local name resolution, you need to configure the name resolution locally in each system.

- We recommend that you use Domain Name Services (DNS) instead of local name resolution, because it offers easy and centralized management of your names and IP addresses.

The order for name resolution is configurable in AIX. The default order can be modified creating the file /etc/netsvc.conf. Both the default and /etc/netsvc.conf can be overwritten defining and exporting the environment variable NSORDER. The default, which is hardcoded, is:

    hosts=bind, nis=auth, local

An example of /etc/netsvc.conf could be:

    hosts=local,bind,nis

Here, /etc/hosts will be queried first, then DNS, and finally NIS (only these three keywords are allowed, after the main keyword *hosts=*). You can also define any of the services as authoritative, which means that name resolution will end in this service, even if it could not resolve the name. Only if an authoritative service is not available will the next service specified be queried. An example of using authoritative service could be:

    hosts=nis=auth,bind,local

The variable NSORDER has the same format as the *hosts* keyword in /etc/netsvc.conf. An example of using the environment variable NSORDER could be:

export NSORDER=nis,local,bin

Additional points of interest about name resolution and Domino 4.6.1

- It is not necessary to match the name of the Domino server with the hostname of the system.

- Once you have your Domino server up and running, clients do not need to resolve the name of the Domino server out of the *Connection Documents* if the connections are configured using the *Advanced* section IP address field. This information could be helpful if you have a problem with name resolution.

3.1.4.4 Environment Setup List

Here is a list of the steps you should take to prepare your AIX system before starting the Domino 4.6.1 installation.

1. Update your software levels if necessary. Refer to 3.1.2, "Operating System Level" on page 27.

2. Create the file system for the Domino binaries (if you decide to place them in a separate file system).

3. Create a file system for the Domino databases.

4. Mount the new file systems.

5. Create the AIX group account notes.

6. Create the AIX user account for ownership of the Domino software.

7. Change the owner and group of the Domino file systems.

8. Prepare the $HOME/.profile of the Domino user.

9. Set up your network configuration.

3.1.5 Installation of the Domino Binaries

Before starting the installation process, if you created a separate file system for the Domino binaries, for example /usr/lpp/lotus, remove the directory lost+found from it (otherwise the final check of the installation process will complain about it). After the installation, you must recreate the lost+found directory under /usr/lpp/lotus.

To avoid this problem, you can just create a file system with the name /usr/domino and install the software in /usr/domino/lotus.

The following steps and screens will take you through the installation of Lotus Domino 4.6.1 on AIX.

1. Log on to the system as user root.

   Change the directory to the base path of the installation software. For example, if you have the installation software in CD-ROM, it can be /cdrom/unix.

2. Type `./install` and press Enter. The Installation Welcome screen will be displayed. Press the Tab key to go to the following screen.

3. Press the Tab key to continue. On the following screen you will be asked to read and accept the Lotus Notes Software Agreement.

4. The default value of this screen is **No**; you should press the Spacebar to change the value to **Yes**. Then press the Tab key to accept the new value and continue to the following screen.

5. The default value here is **Notes Workstation**. To change the value to **Domino Server**, press the Spacebar (if you install the Domino Server, the Workstation software is also installed) and then press the Tab key to continue to the following screen.

6. Now the installation process asks you if you want to install the Domino Advanced Services. The default value for this screen is **No**. As we are going to go to a clustered or partitioned configuration, press the Spacebar to change the value to **Yes**. Then press the Tab key to accept the new value and continue to the following screen.

7. Press the Spacebar to select not to install Solaris software. Then press Tab to continue.

8. Using the Spacebar, choose **Yes** if you want to install the optional online documentation. Press the Tab key to go to the next installation screen.

9. This screen asks whether the Domino server software will be installed on the system that will be running the Domino server. Lotus strongly recommends that the Domino server run on the system where it was installed. Press the Tab key to accept the Lotus recommendation.

10. In this screen you are asked if you want to install to one computer or to more than one computer. Press the Tab key to accept to install in one computer.

    If you want to install the Domino binaries in more than one computer automatically, you need to create the file /tmp/notes_install_hosts, containing all the systems where you want to install the binaries, and configure all these systems to accept automatic rsh requests from the root user of the system where you are running the actual installation process

(by creating the /.rhosts file in all the systems). Use the Spacebar to change the value to [more than one computer] and follow the screens for multi-installation. The installation process will sequentially install all the systems you included in the file /tmp/notes_install_hosts.

11. The screen indicates the hostname of the machine on which we will install the software. The default value is the hostname of the machine where we are installing the software. Press the Tab key to go to the next installation screen or the Enter key if you want to change it.

12. In this screen, you are asked for the directory name where the Domino binaries will be installed. The default value is /opt/lotus. We changed it to /usr/lpp/lotus. Press Enter to change the value to /usr/lpp/lotus, and once you type the new value press Enter and then Tab to accept the changed value and go to the next screen.

13. Specify the installation owner. The default value is notes. You can change it at your convenience. If this server is not going to be a partitioned server, use the name notes for this user.

14. Press the Tab key to go to specify the installation group ID that will own the installed software. Accept the default group notes.

15. Press the Tab key to display the confirmation screen for the actual parameters.

16. Press the Tab key to start the installation. On the following screen the installation process will show you the configuration settings. If they are correct press the Tab key to continue the installation.

17. Press the Tab key to continue after the installation configuration validation or correct any errors and start the installation again.

Now you are ready to run the Notes Workstation to set up the Domino Server.

3.1.6 Setting Up the Domino Server

After you installed the Domino code, you must set up the Domino server. In this process, the server files are copied to the Notes data directory and the default Domino configuration is created in notes.ini file.

Before starting the setup of the Domino Server:

- Be sure the Domino data file system (for example, /notes) is mounted.

- Be sure you changed the operating system parameters as recommended in 3.1.3, "Operating System Parameters" on page 29.

- You must run the setup from a graphical workstation (local graphical display or PC client with X-Windows software correctly configured).

- Set and export the DISPLAY variable.
- Allow write access to $DISPLAY.

Now you are ready to do the following steps to set up your Domino server for cluster or partitioned future configurations:

1. Log on to the system as user notes.
2. Modify the PATH variable in $HOME/.profile as shown in the following example:

   ```
   PATH=$PATH:$HOME/data:/opt/lotus/bin
   ```

3. If you are using CDE, then you need to comment out the last line of the file $HOME/.dtprofile as shown in the following example:

   ```
   DTSOURCEPROFILE=true
   ```

4. Type `./notes`

The following window will appear:
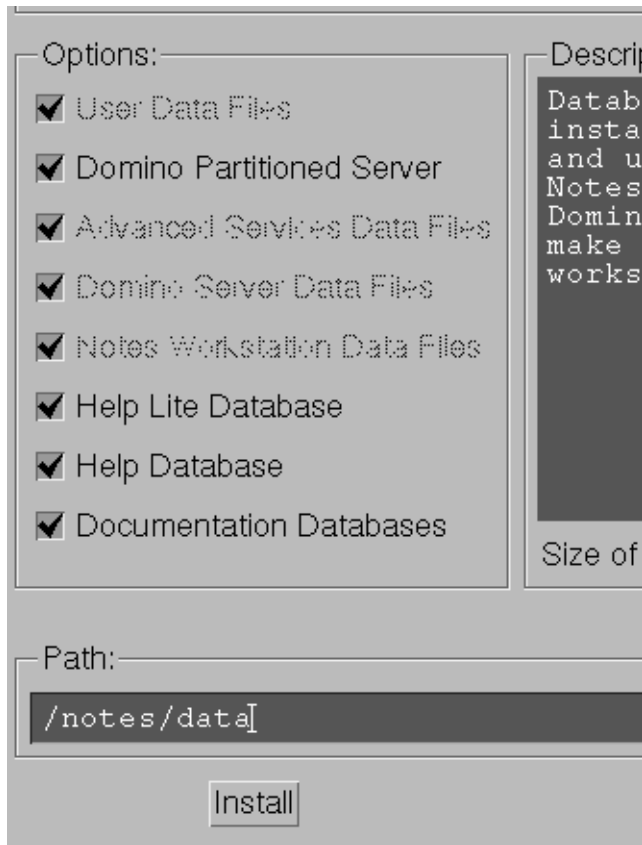
Figure 4. Domino Data Installation Window (Partial View)

5. Select the items that you want to install from the initial setup window, including all the Advanced Services. The Help Lite Database, Help Database and Documentation Databases are up to you. The rest must be selected. We recommend that you set up the data directory on /notes/data instead of /notes/notesr4 because it is more understandable. Once you have selected the items that you want to install, click the **Install** button. After successful installation of the items you selected, the following message will be displayed on the data installation window:
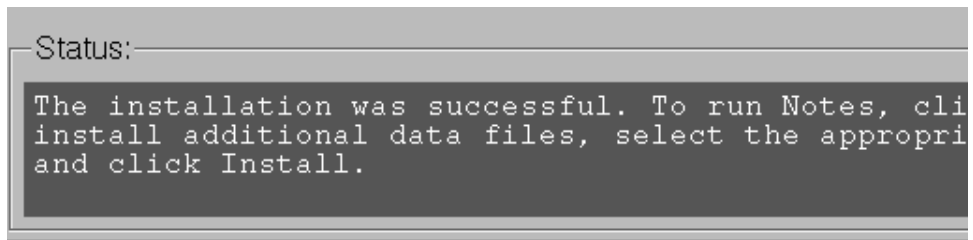
Figure 5. Successful Installation Window (Partial View)

Click the **Quit** button and the Notes client workstation will start giving you the initial Domino server setup window. We show you the Domino server setup windows because they have changed from Domino 4.5.x. If this is the first Domino server of your organization, go to 3.1.6.1, "Setting Up the First Domino Server" on page 37; if it is an additional server, go to 3.1.6.2, "Setting Up an Additional Domino Server" on page 41.

3.1.6.1 Setting Up the First Domino Server

On the Create a New Domino Server window (Figure 6 on page 37), select the **First Domino Server** option and click the **Next** button (**>**).
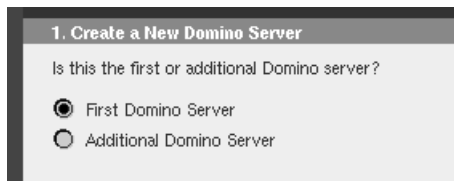
Figure 6. Create a New Domino Server Window (Partial View)

On the following window (Figure 7 on page 37), select **Advanced Setup** and click the **Next** button (**>**).
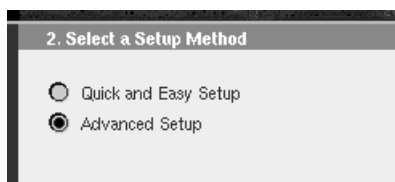
Figure 7. Select a Setup Method Window (Partial View)

On the following window (Figure 8 on page 38), the setup process allows you to select which Domino features you want to configure on the Domino startup. For each option you select, the setup process will add the task to the *ServerTasks* line in the notes.ini file. We recommend that you select none of the Domino subsystems available in this window. It is better to make up the simplest initial Domino configuration, and later, once you are sure that it is working perfectly, configure and start additional Domino features.
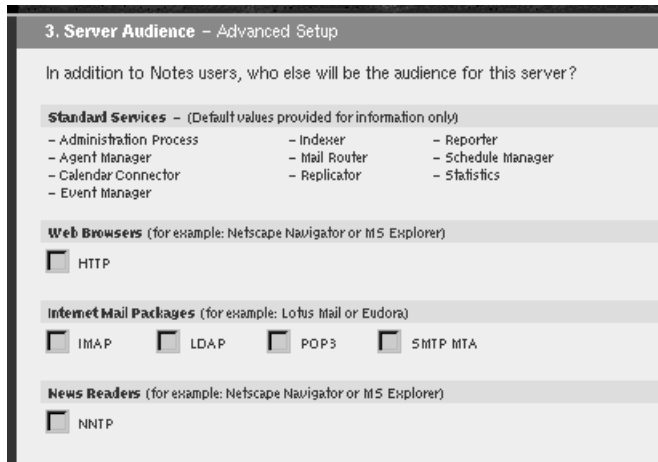
Figure 8. Server Audience - Advanced Setup Window (Partial View)

Once you have finished your selection, click the **Next** button (**>**). The Administration Settings - Advanced Setup window (Figure 9 on page 39) will appear.

Figure 9. Administration Settings - Advanced Setup Window (Partial View)

Fill in the fields at your convenience. Be sure that the name of the Domino server you are setting up in the Server Name field is correctly resolved in your system and in all the clients or servers which will connect with this Domino server. Change the server name to the short TCP/IP hostname, if the hostname was fully qualified. Once you have finished the settings, click the **Finish** button. After some initial processing, the following window (Figure 10 on page 40) will appear:
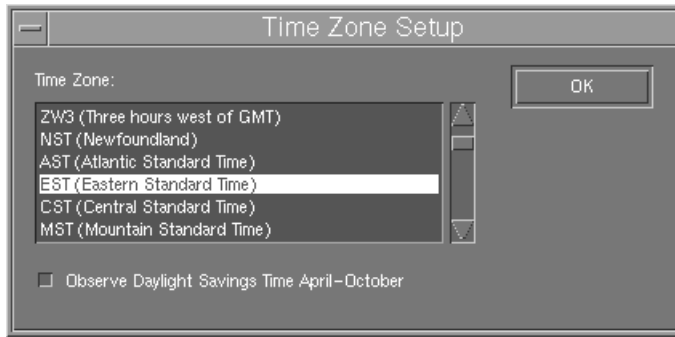
Figure 10. Time Zone Setup Window

Select the appropriate time zone and check Observe Daylight Savings Time April-October at your convenience. Then click the **OK** button and after some processing the following window will be displayed:

Figure 11. Congratulations Window (Partial View)

Click the **Exit to Workspace** button on the Congratulations Window, (Figure 11 on page 40) and close the Domino Setup Notes database to go to the Workspace. Now you can go to section 3.1.6.3, "Additional Setup" on page 44 to do the additional setup steps.

3.1.6.2 Setting Up an Additional Domino Server

If you are installing an additional Domino server in your organization, then you need to create a Notes Server ID on your existing server for the new server you want to install before continuing with the setup.

On the Create a New Domino Server window select Additional Domino Server and click the **Next** button (**>**).
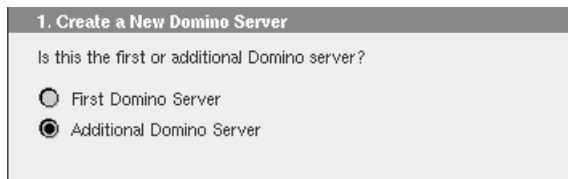
Figure 12. Create a New Domino Server Window (partial view)

On the Select a Setup Method window, select Advanced Setup and click the **Next** button (**>**).
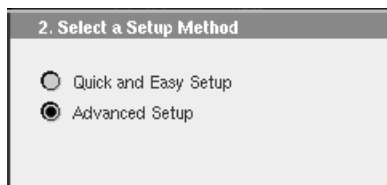
Figure 13. Select a Setup Method Window (Partial View)

On the following window (Figure 14 on page 42), follow the same guidelines we recommended in the previous section 3.1.6.1, "Setting Up the First Domino Server" on page 37 for the same step.
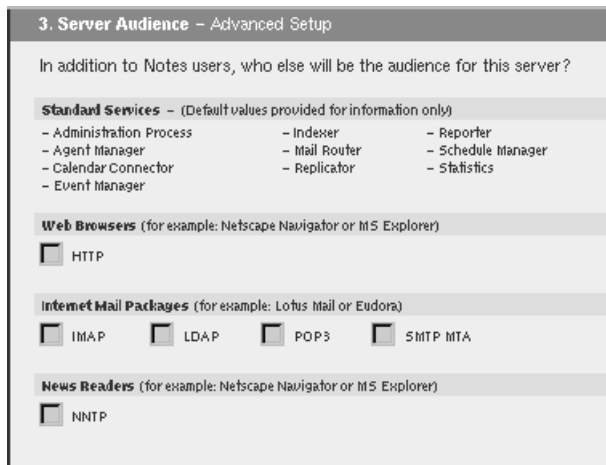
Figure 14. Server Audience - Advanced Setup Window (Partial View)

On the Administration Settings - Advanced Setup window (Figure 15 on page 43), set **Server ID supplied in a file** and in the field Get Address Book from which Server, write the name of your Domino server.

Important

Be sure that the Domino server from which you will get the Name and Address Book is up and running, otherwise the setup process will ask you for a secondary server to get the Address Book from.
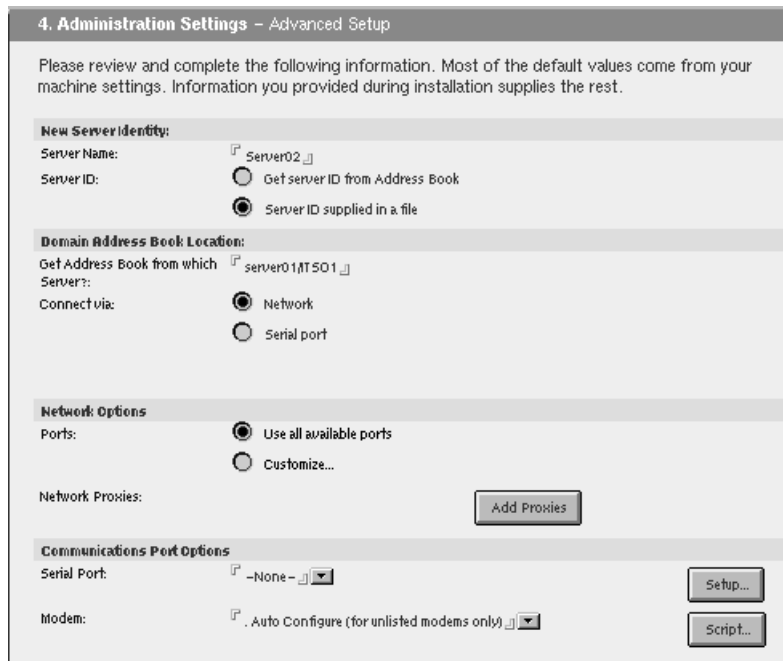
Figure 15. Administration Settings - Advanced Setup (Partial View)

Once you have finished this form, click the **Finish** button. After some processing, the following window will appear:

Figure 16. Enter Server ID File Name Window

Select the ID file you previously created in your primary server and click **OK**. After some processing, the Congratulations window will be displayed.
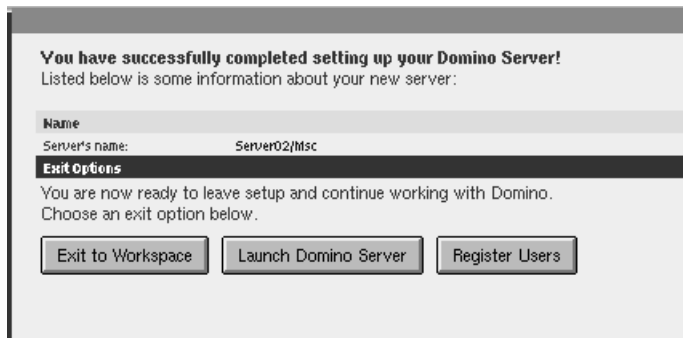
Figure 17. Congratulations Window (Partial View)

On this window, click the **Exit to Workspace** button and close the Domino Setup Notes database to reach the Workspace. Now you can go to section 3.1.6.3, "Additional Setup" on page 44 to do the additional setup steps.

3.1.6.3 Additional Setup

The following setups will prepare your Domino server that any possible future configuration.

We recommend that you do all these steps using the local Notes workstation of the Domino server that you are setting up.

1. We recommend that you set the Administration Server in the Name and Address Book (NAB).

   The Administration process cannot start if this field is not configured. If the Administration process is not properly running, you cannot create a Domino cluster. (Even if you are not planning to use Domino Clustering, it is advisable to have the Administration process correctly up and running)

   After clicking with the right button in the NAB, select Access Control -> Advanced and select the Domino server of your organization that will be the Administration server.

   You only need to do this in the Domino server from which the rest of the Domino servers will get the NAB. Normally, it will be your first Domino server.

2. We recommend that you register the initial people who will be the administrators of the Domino servers.

It is a good idea to create a set of administrator users for your Domino installation at the beginning.

Select File Tools -> Server Administration -> People -> Register Person and register all your administrators.

3. We recommend that you create the Admini9strator group with the users previously created.

This will simplify the management of your Domino installation.

Select Create - Group and define the group Administrators, adding to it all your administrator users.

4. We recommend that you set the Administrators field of the server document to the new group Administrators.

Open the NAB and select Server -> Servers (click on the server) -> Edit, remove the actual value and add the group Administrators.

5. We recommend that you check the Network Configuration section in the server document.

Open the NAB and select Server -> Servers -> (double click on server) -> Network Configuration, check the information of this section.

6. We recommend that you set up the Access Control List (ACL) for the NAB of your server.

This reduces the ACL to the minimum necessary, simplifying your configuration.

Attention

Never remove the Administration Server from the ACL of the NAB. If you do that, it will be automatically removed from the Advanced section and the Administration process will not start. The same applies to the rest of the databases (if you remove the Administration server from the ACL, it is removed as Administration server in the Advanced section), but it is critical for the NAB because that is where the Administration process searches for the Administration server.

You should never make changes to these settings in your secondary Domino servers of the common databases that are replicated between them.

Now click with the right button in the NAB, select Access Control, and in the Basics section, do the following:

- From the current settings, remove all of the servers except the Administration Server, LocalDomainServers and OtherDomainServers.
- Add the Administrators group.
- Set the field User type to person group.
- Set the field Access to Manager.
- Set the capabilities and Roles at your convenience.

You only need to do it in the Domino server from which the rest of the Domino servers will get the NAB. Normally, this will your first Domino server.

7. Create the Certification Log database (certlog.nsf)

This database is not created by default during the installation and initial setup processes. Creating it allows you to log and record all the certification tasks.

Select File - Database - New and in the New Database window:

- Set the field Server to the Domino server we are configuring or leave Local if you are using the local Notes workstation.
- Select the template Certification Log for the new database.
- Set the field Title to certlog (automatically the field File Name is set to certlog.nsf, and it must be the name of the file).
- Click the OK button.

8. We recommend that you set the ACL for the Certification Log (certlog.nsf), Server Certificate Admin (certsrv.nsf) and Notes Log (log.nsf) databases as you did with the NAB database.

9. Start your Domino server.

You should get the following screen:

```
...........
03/04/98 19:12:18       Database Replicator started
03/04/98 19:12:23       Index update process started
03/04/98 19:12:29       Stats agent started
03/04/98 19:12:34       Agent Manager started
03/04/98 19:12:36       JVM: Java Virtual Machine initialized.
03/04/98 19:12:36       AMgr: Executive '1' started
03/04/98 19:12:39       Schedule Manager started
03/04/98 19:12:39       SchedMgr: Validating Schedule Database
03/04/98 19:12:39       SchedMgr: Done validating Schedule Database
03/04/98 19:12:44       Calendar Connector started
03/04/98 19:12:49       Server01/ITSO1 is the Administration Server of
the Public
Address Book.
03/04/98 19:12:49       Creating Administration Requests database.
03/04/98 19:12:49       Administration Process started
03/04/98 19:12:54       Reporter started
...........
```

Figure 18. Server Console Output on the First Start

Once the Administration process is started, it creates the Administration
Requests database.

Check Point

Make sure that the Administration process has started correctly and has
created the Administration Requests database.

10.We recommend that you set the ACL for the Administration Requests
database of your Domino server (admin4.nsf) as you did with the NAB
database.

Now is a good time to register the new Domino servers that you will add soon,
and get their ID files.

After that, you have a Domino configuration perfectly ready to grow with
Domino Advanced Services.

3.1.7 Specifics for SP Systems

If you are going to install a Domino server in an SP system, you should know
some specifics of the software and hardware of the RS/6000 SP.

We strongly recommend that you plan your Domino installation with the system manager of your SP system.

If you are interested in knowing more about the software and hardware of SP systems, refer to the redbook *Inside the SP,* SG24-5145.

The best source of information about how to implement Lotus Domino servers in the RS/6000 SP is the IBM Web site:

```
http://www.rs6000.ibm.com/resource/technology/notesSPcfg
```

Here you will find a lot of useful information about the following:

- Planning the hardware
- Planning SP Domino servers
- Planning SP AIX users
- Planning disk space and configuration
- Planning networks for Domino
- Planning Domino High Availability on the RS/6000 SP
- Domino server installations on the RS/6000 SP
- Domino server configurations on the RS/6000 SP

You will also find information about utilities to do the following:

- Create the AIX users for Notes in parallel
- Create the Notes file systems in parallel
- Create/remove paging spaces in parallel
- Start/stop Domino servers in parallel
- Configure Domino servers to run in the background
- Monitor the Domino servers

Even if you are starting a small Domino implementation on your SP system, follow these recommendations. They will make it easier to grow your Domino environment in your SP system.

3.2 Implementing Domino Partitioned Servers

This section covers the steps for implementing Domino partitioned servers on AIX 4.3 and also the specific steps for SP systems running AIX 4.3/PSSP 2.4.

3.2.1 Lab Environment

We implemented six Domino partitioned servers in our lab using an 8-way SMP RS/6000 with 512MB of RAM, two 2GB disks, and one Ethernet adapter. We assigned unique IP addresses to each partitioned server using IP aliases.

Table 8 shows the configuration we used: the Domino server name, the IP hostname or IP alias, the IP address, the Domino data file system, the AIX user ID and group ID for the Domino server.

Table 8. Domino Partitioned Servers - Domino and IP Configuration

Domino Server Name	IP Hostname or IP Alias	IP Address	Domino Data File System	Domino User ID	Domino Group ID
Server01	server01	192.168.3.1	/notes	notes	notes
Server01p2	server01p2	192.168.3.102	/notes2	notes2	notes
Server01p3	server01p3	192.168.3.103	/notes3	notes3	notes
Server01p4	server01p4	192.168.3.104	/notes4	notes4	notes
Server01p5	server01p5	192.168.3.105	/notes5	notes5	notes
Server01p6	server01p6	192.168.3.106	/notes6	notes6	notes

The Domino organization name we used for the partitioned servers was PARTITIONED.

Figure 19 on page 50 shows a summary of our configuration in the lab.
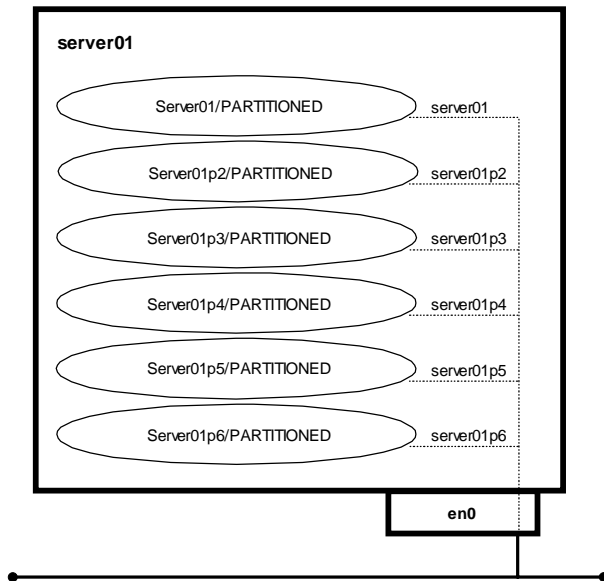
Figure 19. Domino Partitioned Servers - Summary of the Configuration

3.2.2 Installation Steps for AIX

The steps for the installation, registration and setup of the first partitioned server are the same as for a standard non-partitioned Domino server as described in "Installation of Domino 4.6.1 on AIX" on page 25

For the subsequent partitioned servers, only the registration and setup steps are required. No installation is required because all the partitioned servers share the same binaries already installed for the first one.

The installation steps are:

1. Create volume groups, logical volume, and file systems for each partitioned server.

   The following screen shows the configuration we had in the lab.

```
server01:/> df -Ik
Filesystem     1024-blocks      Used      Free %Used Mounted on
/dev/hd4             8192      4268      3924   53% /
/dev/hd2           253952    222828     31124   88% /usr
/dev/hd9var         53248      7940     45308   15% /var
/dev/hd3           24576       916     23660    4% /tmp
/dev/hd1            4096       176      3920    5% /home
/dev/notesbin_lv   184320    179828      4492   98% /usr/lpp/lotus
/dev/notesdat_lv   512000    125252    386748   25% /notes
/dev/notesdat2_lv  122880    110308     12572   90% /notes2
/dev/notesdat3_lv  122880    110308     12572   90% /notes3
/dev/notesdat4_lv  122880    110308     12572   90% /notes4
/dev/notesdat5_lv  122880    110308     12572   90% /notes5
/dev/notesdat6_lv  122880    110308     12572   90% /notes6
```

2. Configure the IP hostname and the IP address for each partitioned server.

   In the lab we configured the IP hostname by adding the servers to the /etc/hosts file.

   To configure the IP aliases for the Ethernet adapter we used the command `chdev -l <adapter> -a alias4=<ipaddress>[,<networkmask>]`, as shown in the following screen.

```
server01:/> chdev -l en0 -a alias4=9.168.3.102,255.255.255.0
server01:/> chdev -l en0 -a alias4=9.168.3.103,255.255.255.0
server01:/> chdev -l en0 -a alias4=9.168.3.104,255.255.255.0
server01:/> chdev -l en0 -a alias4=9.168.3.105,255.255.255.0
server01:/> chdev -l en0 -a alias4=9.168.3.106,255.255.255.0
```

   We used the command `netstat -i` to check that the aliases were added correctly, as shown in the next screen.

```
server01:/> netstat -i
Name Mtu   Network   Address               Ipkts Ierrs   Opkts Oerrs Coll
lo0  16896 link#1                         144666     0  144857     0     0
lo0  16896 127       loopback            144666     0  144857     0     0
lo0  16896 ::1                           144666     0  144857     0     0
en0  1500  link#2    2.60.8c.e8.d2.e1 1078205     0 1063366     0     0
en0  1500  192.168.3 sp2n01             1078205     0 1063366     0     0
en0  1500  192.168.3 server01p2         1078205     0 1063366     0     0
en0  1500  192.168.3 server01p3         1078205     0 1063366     0     0
en0  1500  192.168.3 server01p4         1078205     0 1063366     0     0
en0  1500  192.168.3 server01p5         1078205     0 1063366     0     0
en0  1500  192.168.3 server01p6         1078205     0 1063366     0     0
```

3. Register the Domino partitioned servers.

We registered the partitioned servers as additional servers for the already existing domain called PARTITIONED.

Figure 20 shows the configuration panel we used to register the server ID for Server01p2 with no password.

Figure 20. Basics Registration Panel (Partial View)

Figure 21 on page 52 shows that we are not placing the id file in the Public Address Book, and that the server id file name is server01p2.id.
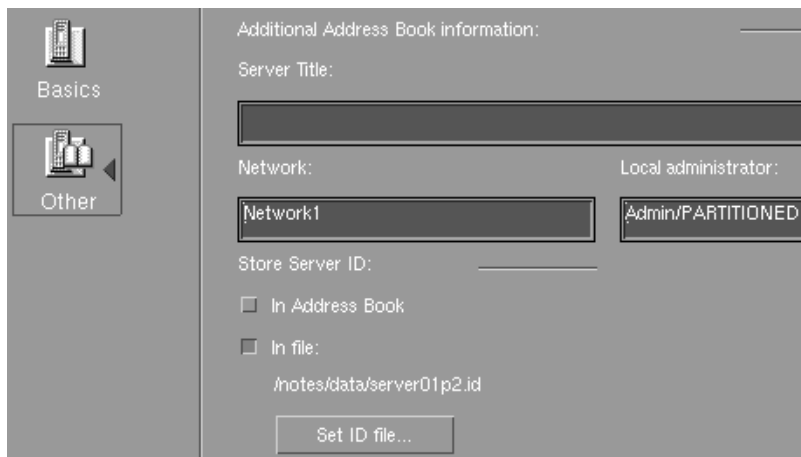
Figure 21. Other Registration Panel (Partial View)

4. Install each partitioned server.

We used /notes2/data as the destination path for installing the Domino data files of the first partitioned server Server01p2; /notes3/data for server Server01p3; and so forth.

Figure 22 on page 53 shows the installation panel that popped up for Server01p2.
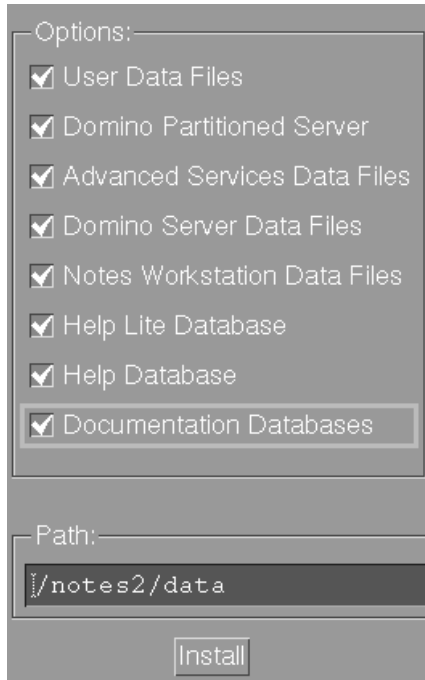
Figure 22. Partitioned Server Installation - Data Files Installation Panel (Partial View)

5. When the installation of the Domino data files ends, the Notes client pops up and opens the Domino setup database.

   Figure 23 on page 54 shows the configuration we used for Server01p2.
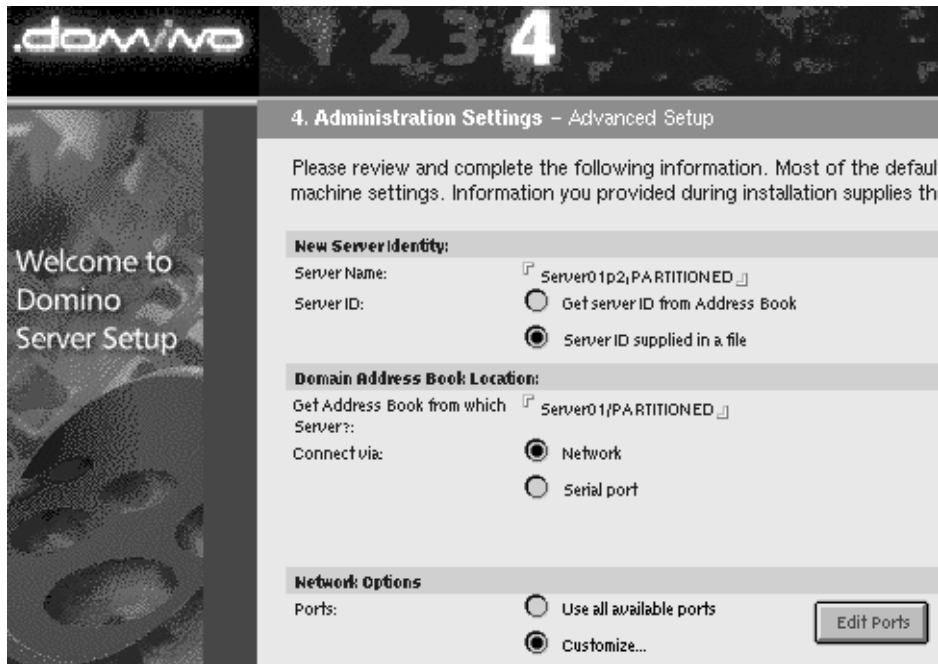
Figure 23. Administration Settings Panel (Partial View)

Also set up the network IP address using the port setup panel.

Figure 24 on page 54 shows the port setup configuration we used.

Figure 24. Port Setup Panel (Partial View)

Each time the setup process is run on a partitioned server, a counter is incremented in the file .sgf.lastpartition, located in the directory /usr/lpp/lotus/notes/latest/common/data/, in order to keep a unique partition number.

Each partitioned server stores its assigned partition number in a file named .sgf.notespartition in its own Domino data directory.

The following screen shows the content of the .sgf.lastpartition file, and the .sgf.notespartition file for the second partitioned server, Server01p2.

```
server01:/> cat /usr/lpp/lotus/notes/latest/common/data/.sgf.lastpartition
6
server01:/> cat /notes2/data/.sgf.notespartition
2
```

6. Edit the Public Address Book to enable the *Bind to host name* and configure the *Host name* field, in the *HTTP Server* section of each partitioned server.

Figure 25 on page 55 shows the configuration we used for Server01p2.

Figure 25. HTTP Server Configuration (Partial View)

7. Edit the notes.ini file for each partitioned server installed, and add the stanzas required for binding the IP address or IP alias to each partitioned server task.

For the Domino server use the following stanza:

`TCPIP_TcpIpAddress=0,<TcpIpAddress>:1352`

For POP3 use the following stanza:

`POP3Address=<TcpIpAddress>`

For IMAP use the following stanza:

`IMAPAddress=<TcpIpAddress>`

For NNTP use the following stanza:

```
NNTPAddress=<TcpIpAddress>
```

For LDAP use the following stanza:

```
LDAPAddress=<TcpIpAddress>
```

The following screen shows the configuration for the first partitioned server.

```
...
TCPIP_TcpIpAddress=0,192.168.3.1:1352
POP3Address=192.168.3.1
IMAPAddress=192.168.3.1
NNTPAddress=192.168.3.1
LDAPAddress=192.168.3.1
...
```

8. Finally, open a session (login) using each Domino AIX user and start the partitioned servers.

   To check each server, use the command `show server` at the console of each partitioned server.

   The following screen shows the output for the server Server01p2.

```
> sh server

Lotus Domino © Server (Release 4.6.1 (Intl) for UNIX) 02/04/98 14:38:10

Server name:          Server01p2/PARTITIONED
Server directory:     /notes2/data
Partition number:     2
Elapsed time:         00:01:01
Transactions/minute:  Last minute: 0; Last hour: 0; Peak: 0
Peak # of sessions:   0 at
Transactions:         0
Shared mail:          Not enabled
Pending mail:  0      Dead mail:  0
>
```

3.2.3 Specifics for SP Systems

Special care has to be taken in the creation of AIX user IDs for Domino partitioned servers in order to have proper home directories. As already explained in "Specifics for SP Systems" on page 47, plan these activities with the system manager of your SP system.

3.3 Implementation of Domino Clustering

This section describes how to install a Lotus Domino cluster. Before you read this section, you should know how to install Domino Enterprise Server as described in "Operating System Parameters" on page 29.

3.3.1 Cluster Installation Prerequisites

All cluster members do not have to run the same version of Domino and Domino clustering, or Domino Enterprise Server. Clusters are available for R4 servers only. Only R4 workstations can take advantage of the cluster failover feature. R4 workstations do not need a special license to use clusters.

All cluster members must use the same set of network protocols.

All cluster members must be connected using a high-speed LAN connection, or a WAN T1 (1.5 Mbit/sec) or better.

All cluster members replicate with all cluster members; there are no masters in the replication scheme.

All cluster members must be in the same Domino domain and share a common Public Address Book.

The Address Book for the Domino domain must have a specified administration server. If you do not specify one, the Administration Process cannot change cluster membership. The server that is the administration server of the Public Address Book does not need to be a member of a cluster or be running the Enterprise Server license. It must, however, be running Domino Release 4.5 or later.

Each server is a member of only one cluster at a time.

Each server must have adequate disk storage space to function as a cluster member. Because clusters usually require more database replicas, servers in clusters require more disk storage space than unclustered servers.

Each server must have adequate processor and memory capacity. In general, clustered servers require more computer power than unclustered servers.

3.3.2 Setting Up and Configuring a Cluster

In addition to the regular setup tasks for a Domino server, there are additional administration tasks for clusters.

Planning a cluster topology includes considering the overall goal for the cluster, database distribution, user work patterns, equipment reliability, cluster topology and capacity requirements. You must also include the network and connectivity requirements of your organization or service, and whether you use a Web Navigator server in your cluster to connect to the Internet.

You must purchase and install the Lotus Domino Enterprise Server license on each server in the cluster.

3.3.2.1 Configuration

You configure a cluster by adding or removing servers to the cluster. Use the *Add to Cluster* or *Remove from Cluster* button in the Public Address Book, as shown in Figure 26.
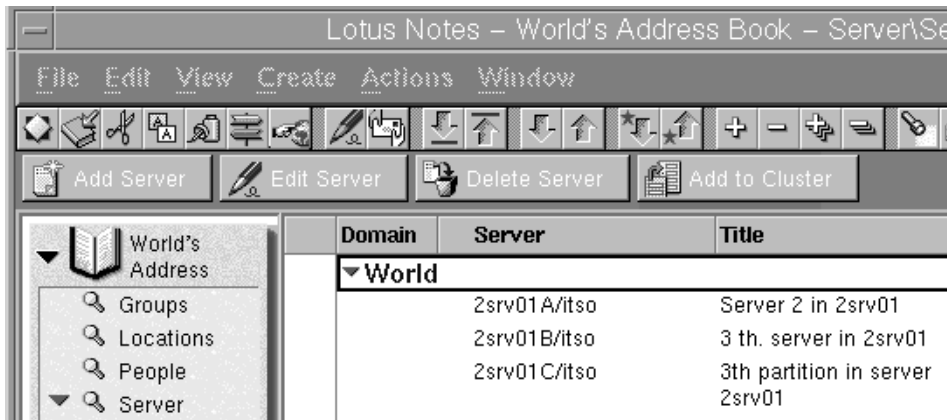
Figure 26. Adding Servers to a Cluster

It is faster to make these changes in the Public Address Book on the administration server, because there you do not have to wait for the next scheduled replication to get your administration request processed.

Before changes in the cluster become effective, the Public Address Book and Administration databases must be replicated between cluster members. You can force this replication, or wait for the next scheduled replication.

The events that take place when a server is added to a cluster, as shown in Figure 27 on page 59, are as follows:

- When you submit the request for adding (or removing) a server to a cluster, a request document is added to the Administration request

database. At the next scheduled replication, this request is replicated to the administration server.

- The Administration server responds to the request by adding (or removing) the server into the cluster definition in the Public Address Book in the Administration server.

- At the next replication the Public Address Book, with the updated cluster information, is replicated to the server. The server then detects the change in the Public Address Book, and starts (or stops) the Cluster Service accordingly.
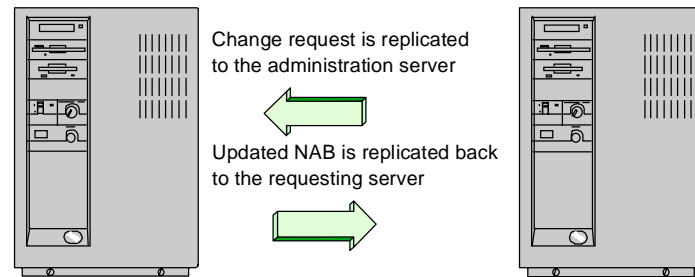
Figure 27. Processing Cluster Update Request

Each server in a cluster must use the same set of network protocols, and the servers should be connected to the same physical LAN. In case of a failure in one of the servers, the others should take over. You can also optionally create a private network restricted to cluster traffic only.

3.3.2.2 Databases and Replicas
If you are installing clustering for availability purposes, you would need two or three database copies; for load balancing purposes, you would need more copies, depending on your configuration.

3.3.3 Cluster Components
Each server in a cluster runs cluster components, which are installed with the Lotus Domino Enterprise Server license. These components, with the help of the Administration Process, perform the cluster management and monitoring tasks that let you administer a cluster.

3.3.3.1 Cluster Manager
The Cluster Manager tracks the state of all members in a cluster. It keeps a list of which servers in the cluster are currently available and maintains information about the workload on each server. You can view this information

by typing the `show cluster` command at the server console, as shown in the following screen.

```
> sh cluster
Cluster Information
 Cluster name: Lyngby, Server name: RISC73/ITSOMSC
 Server cluster probe timeout: 1 minute(s)
 Server cluster probe count: 15811
 Server availability threshold: 0
 Server availability index: 100 (state: AVAILABLE)
 Cluster members (1)...
 Server: RISC73/ITSOMSC, availability index: 100
```

The Cluster Manager resides on each server in a cluster. When you add a server to a cluster, Domino automatically starts the Cluster Manager on that server. A Cluster Manager is responsible for tracking which servers belong to its cluster. To monitor the cluster, each Cluster Manager exchanges periodic messages, called *probes*, with other servers. It knows which servers in the cluster are currently available and what the workload is on each server. The manager determines which cluster servers have a replica of a requested database and redirects the request to the appropriate server. The Cluster Manager uses the Cluster Database Directory to select the servers and directs the open database request accordingly.

The tasks of the Cluster Manager include:

- Determining which servers belong to the cluster by periodically monitoring the Public Address Book for changes to the ClusterName field in the server record or to the cluster membership list.

- Monitoring current server availability and workload in the cluster.

- Advising other Cluster Managers of changes in cluster server status.

- Redirecting database requests based on known cluster server status.

- Balancing server workload in the cluster based on known cluster server status.

- Logging failover and workload balance events in the server log file.

The first task of a new Cluster Manager is to check the Public Address Book to determine which servers belong to the cluster. This information is maintained in memory and used by the Cluster Manager to exchange periodic messages with other Cluster Managers. This in-memory list is called the *cluster cache*.

Through these exchanges, Cluster Managers are able to track the availability and workload of servers in the cluster and to perform the functions previously mentioned.

The Cluster Manager triggers cluster events such as failover and workload balancing when there is a change in the status of the cluster. And the cluster manager provides statistics and logging.

3.3.3.2 Cluster Replicator

The *Cluster Replicator* (CLREPL) task is responsible for the tight synchronization of data among databases and their replicas in a cluster. Unlike the normal scheduled replication set up by use of connection documents between the servers, the cluster replication is event driven.

The cluster replication task runs on each cluster server and continuously responds to changes to its servers' databases by pushing the changes to each of the other replicas within the cluster. It is this continuous transfer of changes that allows Domino to achieve closely synchronized databases. The cluster replicator also replicates private folders that are stored in the databases.

Note

More than one cluster replicator can be running at the same time, and this may improve performance in clusters with many servers and heavy cluster traffic. To start more cluster replicators, add more occurrences of CLREPL on the ServerTasks line in notes.ini.

The cluster replicator uses the *Cluster Database Directory* (cldbdir.nsf) to determine which local databases have replicas on other cluster members. This information determines which modifications trigger replication events, and which replicas receive the changes.

The cluster replicator stores this information in memory. When the cluster replicator detects changes in the cluster database directory, it updates the in-memory information; for example, adding or deleting a database or disabling replication for a database.

The cluster replicator task only maintains replications to other servers in the cluster; the standard replicator task (REPLICA) is still responsible for replicating changes to and from servers outside of the cluster.

The database list in cldbdir.nsf contains all databases in the cluster, and by default cluster replication is enabled for all database replicas.

Cluster replication occurs only when data changes. Therefore, if databases are not changed, it should not require any performance to keep cluster replication enabled even for databases such as the *Lotus Help* database, help.nsf. We have not been able to measure any load related to cluster replication-enabled read-only databases. However, one could prefer to disable cluster replication for databases where it is not needed.

3.3.3.3 Cluster Replicator Synchronization

When the cluster replicator task is unable to replicate database changes, it retains the information in memory and tries to replicate the changes until it is successful.

To conserve resources that might be wasted on unsuccessful replication attempts and to increase the probability of achieving successful replication on the next retry, the cluster replicator increases the interval between retry attempts from one minute up to one hour for each failed replication.

For example, if a server in the cluster is shut down and therefore not reachable by the cluster replicator on another server, the cluster replicator first retries the replication after waiting for one minute. If this retry attempt fails, the cluster replicator waits an additional two minutes before the next retry. If this attempt fails, the cluster replicator increases the retry delay to four minutes, then eight minutes, and so on up to a maximum delay of one hour.

When the destination server eventually restarts and replication succeeds, all modifications performed on the source database that have not yet been replicated are pushed to the destination database. The cluster replicator brings the destination database into synchronization with the source database regardless of how many changes occurred on the source while the destination database was unavailable.

When the cluster replicator logs replication events, any databases that are awaiting a retry are recorded in the replication log. This allows you to see which databases are not currently in synchronization within the cluster and to see the errors that are preventing replication. After the errors have been corrected and successful replication is completed, the error information is no longer recorded.

3.3.3.4 Cluster Administration Process

The *Cluster Administration Process* (CLADMIN) is responsible for the correct operation of all cluster components. On clustered servers, the process runs automatically at server start-up and whenever the cluster membership changes.

3.3.3.5 Cluster Database Directory

The *Cluster Database Directory* (cldbdir.nsf) database has a replica on each server in the cluster. The database contains information about all the databases and replicas within the cluster. For each database it holds information on replication enabled or disabled; and whether a database is marked in service, out of service, or pending delete.

If you want to disable or enable cluster replication for a database, this is the place to do it. Unfortunately, you have to edit one document for each database you want to enable/disable. You can easily create an enable/disable agent to do the job for you.

3.3.3.6 Cluster Database Directory Manager

The *Cluster Database Directory Manager* (CLDBDIR) task keeps the *Cluster Database Directory* database (cldbdir.nsf) up to date with the most current database information. It also manages databases with cluster-specific attributes such as databases marked out of service or pending delete.

All servers within a cluster run a cluster database directory manager task that creates and maintains the cluster database directory. When you add or remove a database on a server, the CLDBDIR task immediately changes the information in the cluster database directory to show the addition or deletion of that database.

All members of a cluster share a common cluster database directory. When one server updates its list of databases, the cluster replicator replicates the changes to the other servers in the cluster. In this way each cluster member has an up-to-date directory of all databases in the cluster.

3.3.3.7 Workload Balancing

Clusters have a workload balance feature that lets you distribute the workload of heavily-used databases across multiple servers in the cluster. To distribute

the workload, you limit or restrict the work that a server can perform using the following notes.ini settings.

Table 9. Notes.ini Parameters for Workload Balancing

Notes.ini setting	Description	Server state	Use
Server_Availability _Threshold	Specifies the level of system resources available on the server	BUSY	To determine at what workload level users will be redirected to another server
Server_MaxUsers	Sets the maximum number of users allowed to access the server at the same time	MAXUSERS	To determine at what usage level new users will be redirected to another server
Server_Restricted	Specifies whether user access to the server is enabled or disabled	RESTRICTED	To prevent users from accessing databases on a specific server

If the workload on your cluster is normally well distributed, consider configuring the cluster for failover only, and not workload balancing.

3.3.3.8 Failover configuration

Only R4 workstations can failover. Notes R3 workstations that try to open a database on a server that is unreachable receive the message `Server is not responding`. R3 workstations are allowed to access a busy server.

If the workstation is running R4 and tries to open a database on an unavailable server, the request fails over to a replica of the database on another server in the cluster if possible. If not possible, a busy server allows the workstation access. Maxusers and restricted servers generate messages (see Table 9 on page 64).

When failover occurs on a server, Domino generates a failover or workload balance event in the log file. If a user selects the icon stack option, the icon of the failover database is added to the top of the icon stack. Otherwise, Notes adds the icon to the workspace, if it does not already exist.

3.3.4 Clustering Definitions in the notes.ini File

You can have two kinds of clustering definitions in the notes.ini file:

- Definitions for clustering tasks
- Definitions for clustering settings

3.3.4.1 Clustering Tasks in the notes.ini File

When running a cluster, some server tasks are added to the *ServerTask* stanza in the notes.ini file:

- **CLADMIN** - The cluster administration process checks the operation of all cluster components.

- **CLDBDIR** - The cluster database directory manager updates the cluster database directory and manages databases with cluster-specific attributes.

- **CLREPL** - The cluster replicator performs the database replication within a cluster.

CLDBDIR and CLREPL are added automatically. CLADMIN is not added and does not need to be. It is started through an internal triggering mechanism based on the cluster name in server document.

3.3.4.2 Clustering Settings in the Notes.ini File

The following settings can be used for clustering in the notes.ini file:

- MailClusterFailover - Enables or disables mail router request failover. If users have replicas of mail files located on multiple servers, you can set this variable in the notes.ini file of all R4 servers in the domain to let users receive mail from servers within and outside the cluster when their home servers are down. This parameter lets the mail router deliver the mail to a replica of the mail database.

  Syntax: `MailClusterFailover=value`

  Use the following values to set this variable:

  - 0 - Disables mail router request failover.
  - 1 - Enables mail router request failover.

- RTR_Logging - Controls monitoring of Cluster Replicator activity.

  Syntax: `RTR_Logging=value`

  Use the following values to set this variable:

  - 0 - Disables monitoring of the cluster replicator.
  - 1 - Enables monitoring of the cluster replicator.

- Server_Availability_Threshold - Specifies the server workload level at which Domino attempts to redirect new user requests to another server in the cluster and the server attains a busy state.

  By setting this value for each server in a cluster, you determine how the workload is distributed among cluster members. Valid values are 0 to 100.

Domino compares this value against a server's availability index; when the availability index falls bellow the Server_Availability_Threshold value, the server becomes busy. A Server_Availability_Threshold value of zero (0) indicates a fully available state and workload balancing is disabled; a value of 100 indicates the server is busy (since the availability index can never be greater than 100) and the cluster manager then tries to redirect user requests to more available cluster members.

Syntax: `Server_Availability_Threshold=value`

Use the following values to set this variable:

- 0 - Default value; disables workload balancing.
- 100 - Enables workload balancing and redirects all requests.

- Server_Cluster_Default_Port - Lets you specify the port used for intracluster network traffic.

Syntax: `Server_Cluster_Default_Port=port`

Use the following value to set this variable:

- `port` - any defined LAN port; for details refer to "Cluster Network Communication" on page 71.

- Server_MaxUsers - Sets the maximum number of users allowed to access a server. When this number is reached, the server state becomes MAXUSERS, and the server stops accepting new Database Open requests.

Syntax: `Server_MaxUsers=number_of_users`

Use the following values to set this variable:

- 0 - Unlimited access to server by users.
- any number - Restricts number of active users to the number you specify.

- Server_Restricted - Specifies whether access to the server is restricted. Enables or disables access to a server. If access is disabled, the server does not accept new Open Database requests.

Syntax: `Server_Restricted=value`

Use the following values to set this variable:

- 0 - Server access is unrestricted.
- 1 - Server access is restricted for the current server session. Restarting the server clears the setting in notes.ini file.
- 2 - Server access is restricted persistently, even after server restarts.

3.3.5 Creating the First Server in a Cluster

Before you start setting up a cluster, check that the administration server runs on all servers in your domain, or at least on those servers you will include in the cluster.

The easiest way to check the function of your administration server is to:

- Open the Public Address Book for each server in your domain and check that the server build number is filled out.

  This ensures that the administration server is or has been running on the server.

- Check the administration server in the Public Address Book.

  This is done by selecting, with the right mouse button, the Public Address Book at the Notes client workspace. Then select **Access Control...** option, and in the Advanced panel, set or confirm the administration server is correct.

  This field should be filled in with the same server name in all replicas of the Public Address Book in the domain.

- Check the Administration Requests database (admin4.nsf), to see if all requests are carried out.

  Errors or pending request indicate that the Cluster creation is likely to fail.

- At the server console at each server, issue the command `Show Task,` and check that the Adminp process is running.

Now when all checks are done, it is easy to create the cluster and add nodes to it, as follows:

1. Open the Public Address Book in the server view.

2. Mark one or more servers and press the **Add server to Cluster** button, as shown in Figure 28.
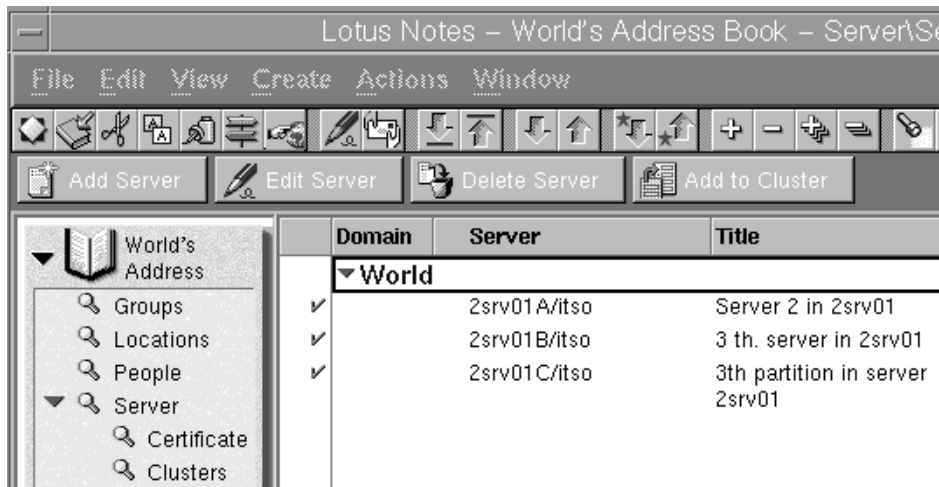
Figure 28. Adding a Server to a Cluster - Add and Delete Button in the PAB

3. When you are asked for confirmation, press **Yes** to confirm, as shown in Figure 29.
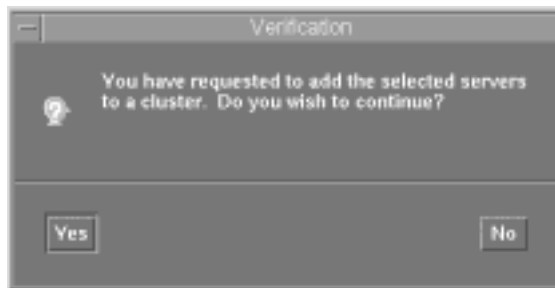
Figure 29. Adding a Server to a Cluster - Confirmation Panel

4. Select an existing cluster name or select **Create New Cluster** to create a new cluster, as shown in Figure 30.
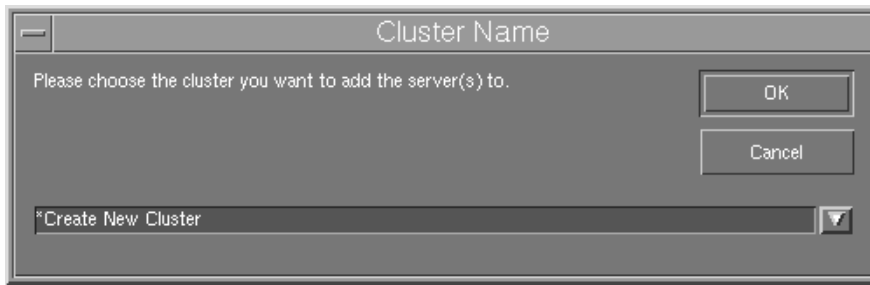
Figure 30. Adding a Server to a Cluster - Cluster Selection Panel

5. If you select to create a new cluster, type the name of the new cluster in the New Cluster Name panel, as shown in Figure 31 on page 69.

Figure 31. Adding a Server to a Cluster - New Cluster Name Panel

After you press **OK**, the request for adding a server to the cluster is placed in the Administration database, as shown in Figure 32.
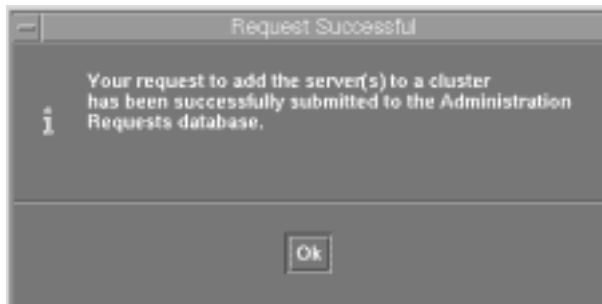
Figure 32. Adding a Server to a Cluster - Request Successfully Submitted Panel

6. Open the Administration database, and check that the request for adding one or more databases to a cluster was submitted and granted, as shown in Figure 33 on page 70.
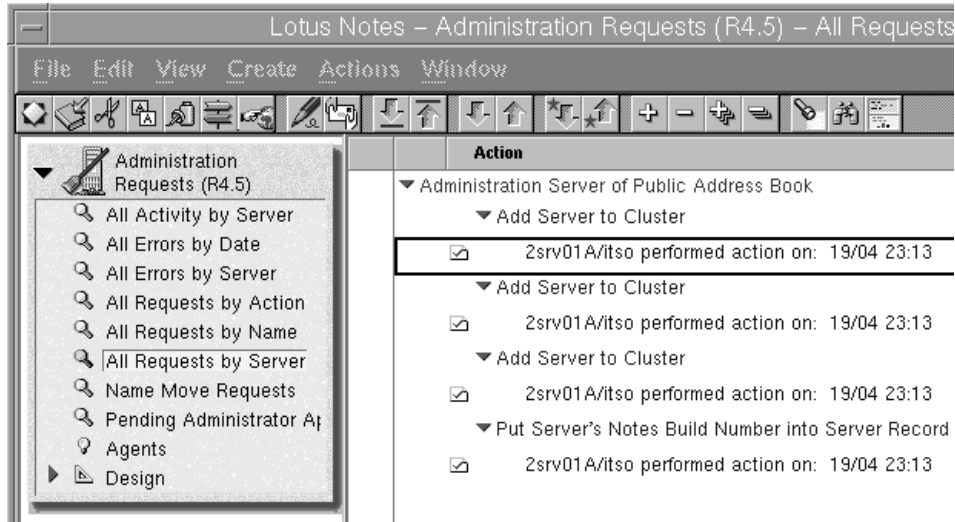
Figure 33. Check Administration Requests (Partial View)

Each server in the cluster starts the cluster service and begins logging information to the console and into the log.nsf database.

In the notes.ini file, the cldbdir and clrepl tasks are added to the ServerTasks stanza, as well as to two new stanzas. The following screen shows these modifications in the notes.ini file.

```
...
ServerTasks=Router,Replica,Update,Amgr,Sched,AdminP,Report,Event,Billing,Cldbdir,Clrepl
...
SERVER_CLUSTER_ON=1
MailClusterFailover=1
...
```

Also, a cluster directory database is created with a replica copy on each server in the cluster. This database contains a list of all databases on each server.

Since this database on each server only holds the local file names, no cluster replication occurs until it is replicated with the normal replication process,

either manually from the server console or by configuring connection documents and waiting for the next scheduled replication.

After the first replication, each server in the cluster holds a copy of the cluster database directory database populated with all databases on all the servers in the cluster. By default, the cluster replication is enabled for all databases.

Whenever you create or delete a database on a server, the cluster manager detects the change, and the cluster database directory database is updated accordingly. Since cluster replication is enabled for the cluster database directory database, any changes are replicated immediately and the cluster database directory remains synchronized across the cluster.

3.3.6 Cluster Network Communication

Cluster networks support two types of communication:

- Between cluster servers and workstations
- Among cluster servers

3.3.6.1 Communication between Cluster Servers and Workstations

For communication between cluster servers and workstations, you must configure all servers to support the same set of network protocols. This configuration enables any cluster member to handle any user request for a database that users can access.

3.3.6.2 Communication among Cluster Servers

For communication among cluster servers, you can optionally configure servers to use a private local area network (LAN). A private LAN for intra-cluster communication can improve cluster performance.

To configure a private LAN for cluster communication on TCP/IP, you must first define two or more LAN adapters in your operating system. The adapters should be in different IP subnets.

For example, in our lab, server BMW uses LAN1=9.12.1.10 for workstation communication and LAN2=9.12.2.10 for cluster communication.

To bind workstation traffic to a specific LAN adapter and the cluster traffic to another, define two Domino port names and assign network addresses to them.

To define a Domino IP port name, use the following stanza in the notes.ini file:

```
<DominoPort>=TCP,0,15,0,,12288
```

To bind an IP address to a Domino IP port, use the following stanza in the notes.ini file:

```
<DominoPort>_TcpIpAddress=0,<IPaddress>:<IPportnumber>
```

For example, the configuration for server BMW had defined a port WSSERVICE for workstation service, and port TCPCLUSTER for cluster traffic. As shown in the following screen, these ports were bound to different IP addresses: WSSERVICE was bound to 192.12.1.10 using port 1352; and TCPCLUSTER was bound to 192.12.2.20 also using port 1352. The cluster replication port was bound to the TCPCLUSTER port.

```
...
WSSERVICE=TCP, 0, 15, 0,,12288
TCPCLUSTER=TCP, 0, 15, 0,,12288
WSSERVICE_TcpIpAddress=0,192.12.1.10:1352
TCPCLUSTER_TcpIpAddress=0,192.12.2.10:1352
Server_Cluster_Default_Port=TCPCLUSTER
...
```

Finally, you must define both ports for server traffic in the Public Address Book as shown in Figure 34.

Figure 34. Cluster Ports Defined in notes.ini, not as Domino General Ports

At the server console, you can check how the server uses the two or more IP LANs with the show port tcpip command.

3.4 Implementation of HACMP for AIX

HACMP for AIX allows you to implement a highly available Domino configuration in several ways. In this section we describe how you can configure HACMP for AIX to achieve it.

To understand our explanations, you should be familiar with HACMP.

To plan your highly available configuration, refer to 2.2, "High Availability" on page 13.

About HACMP for AIX:

- HACMP for AIX can be configured in up to eight node clusters.
- HACMP for AIX provides High Availability (HA) at the system level.
- HACMP for AIX does not provide HA at the application level. Do not get confused when an HACMP node takes over the node which failed. It also takes over the application services of the failing node, because it detected a failure at the system level. If your Domino server crashes, but there is no problem at the system level, HACMP will not take over the Domino server.

3.4.1 HACMP Configurations

HACMP for AIX can be configured in three ways:

- Standby configurations

  These are redundant hardware configurations where one or more standby nodes stand idle, waiting for a server node failure. If one node fails, one standby node takes over the IP address, the disks, and restarts the Domino server. The standby nodes can perform other tasks (system administration, backup/restore), so when the standby node takes over a failed node, it can stop its main tasks and only run the Domino server until the failed node is repaired.

- Takeover configurations without concurrent access

  In these configurations, all the cluster nodes are doing useful work. Takeover configurations use hardware resources more efficiently than standby configurations, since there are no idle processors. With takeover configurations, performance can degrade after node detachment, because the load on remaining nodes increases. With these configurations, if a node fails, the node that takes over the resources of the failed one is running two partitioned Domino servers. The performance of both Domino servers will be affected, depending on the hardware configuration of the takeover node.

- Takeover configurations with concurrent access

  These configurations use hardware efficiently as well, and minimize service interruption during failover, because the takeover node does not need to acquire the resources released by the failed node. The takeover node already shares ownership of the resources.

Important

This configuration is not supported by Domino servers.

For more information about HA configurations, refer to *High Availability Cluster Multi-Processing for AIX Concepts and Facilities*, SC23-1938 and *High Availability Cluster Multi-Processing for AIX Planning Guide*, SC23-1939.

With HACMP for AIX, you can have up to an eight-node mutual takeover configuration, where each node of the cluster can run a different partitioned Domino server and is able to take over failed nodes, providing the service of the failed partitioned Domino server.

For the client gateway systems, a mutual takeover configuration using cascading resource groups is recommended. In this configuration, both gateway systems are running and each can take over all the services of the other.

3.4.2 Example of HA Domino Configuration Using HACMP for AIX

For installing and configuring an HA configuration using HACMP for AIX, refer to:

- *High Availability Cluster Multi-Processing for AIX Installation Guide*, SC23-1940
- *High Availability Cluster Multi-Processing for AIX Administration Guide*, SC23-1941
- *An HACMP Cookbook*, SG24-4553

The HA Domino configuration using HACMP for AIX, which we describe next, is a Two-Node Mutual Takeover node using cascading resource groups. Figure 35 on page 75 shows how we implemented it.

Cars HACMP Cluster
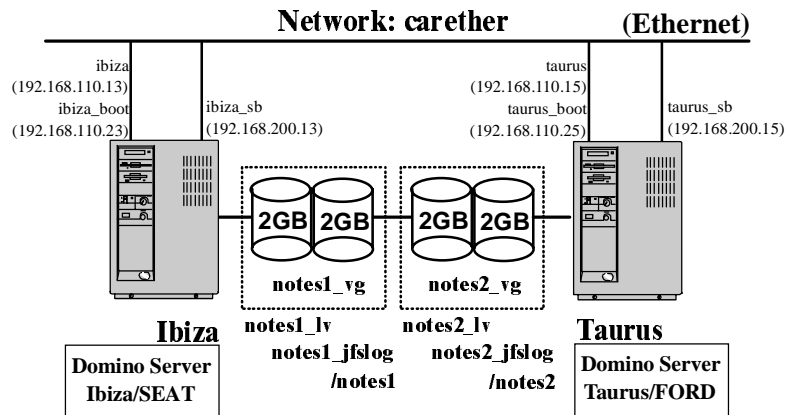
Figure 35. Mutual Takeover Configuration

This configuration consists in two partitioned Domino servers which, on a regular basis, are running in a different system, but in this case, they are installed and configured as if they were two partitioned Domino servers running in the same AIX system.

Hardware configuration:

- Two RS/6000 model 570
- Four external 9333 disks
- Two additional Ethernet adapters in each system for HACMP

Software configuration:

- AIX 4.3.0
- HACMP for AIX 4.2.2
- Domino 4.6.1 with Domino Advanced Services
- Domino binaries installed in the internal disk of each system
- Domino data installed in a separate file system in external 9333 disks
- Each Domino data file system in a separate volume group (no concurrent access)

This configuration is appropriate when each system in the cluster is running critical applications that need to be highly available. To avoid performance impact, each processor should be able to handle the load of more than one system.

Note

As we said before, you can expand this HA configuration to up to an eight-system cluster. Also, you can configure your systems in pairs following this example. In fact, you have multiple choices for configuring your HA Domino environment.

In our example, our requirements were the following:

- Maximum use of the hardware
- High availability for both Domino servers

Here we give you the steps we used to implement the HA configuration of our example. We explain some of the values we set in our configuration. For more details, see the books we suggested. You can follow this guide, changing the parameters according to your installation, to build a successful HA configuration for your Domino servers.

These are the steps:

1. Install the HACMP for AIX software in both systems.
2. Install the Domino binaries on the internal disk of each system.

   Refer to 3.1.5, "Installation of the Domino Binaries" on page 32.
3. Create group notes in both systems with the same group ID (GID).
4. Create users notes1 and notes2 in both systems with the same user ID (UID).
5. Create volume group notes1_vg for Domino data in ibiza with the following initial settings:

   - Activate volume group automatically at system restart, `no`
   - Activate volume group after it is created, `yes`
6. Create volume group notes2_vg for Domino data in taurus with the following initial settings:

   - Activate volume group automatically at system restart, `no`
   - Activate volume group after it is created, `yes`

> **Important**
>
> The name of the shared volume groups for the Domino databases in ibiza and taurus must be different.

These two volume groups will be in different resource groups and will be taken over by the surviving system in case of failure.

7. Create logical volume notes1_lv on ibiza.

8. Create logical volume notes2_lv on taurus.

> **Important**
>
> The name of the shared logical volumes in ibiza and taurus must be different.

9. Create file system /notes1on top of logical volume notes1_lv in ibiza, with the parameter Mount automatically at system restart set to `no`.

10. Create file system /notes2 on top of logical volume notes2_lv in taurus, with the parameter Mount automatically at system restart set to `no`.

11. Rename the journal file system log (jfslog) of each shared volume group in ibiza and taurus, making them different, using `smitty chlv` and modifying the /etc/filesystems file manually.

In our example, the name of the journal file system log LV was lv00 in both systems. To change their names we ran the following commands:

In ibiza: `chlv -n'notes1_jfslog' lv00`

In taurus: `chlv -n'notes2_jfslog' lv00`

After this change we edited the /etc/filesystems file in both systems to modify the entries for the file systems /notes1 (ibiza) and /notes2 (taurus), as follows:

In ibiza:

```
/notes1:
        dev             = /dev/notes1_lv
        vfs             = jfs
        log             = /dev/notes1_jfslog
        mount           = false
        check           = false
        options         = rw
        account         = false
```

In taurus:

```
/notes2:
        dev             = /dev/notes2_lv
        vfs             = jfs
        log             = /dev/notes2_jfslog
        mount           = false
        check           = false
        options         = rw
        account         = false
```

12. Mount the file system /notes1 in ibiza and, as root, issue the command
    `chown notes1.notes /notes1`.

13. Mount the file system /notes2 in taurus and, as root, issue the command
    `chown notes2.notes /notes2`.

14. Log in to ibiza as user notes1 and prepare the $HOME/profile file for the
    user. In the PATH statement you should add the following two directories:

    `export PATH=/opt/lotus/bin:/notes1/data:$PATH`

    (/notes1/data is the Notes data directory.)

15. Log in to taurus as user notes2 and prepare the $HOME/.profile file for the
    user. In the PATH statement you should add the following two directories:

    `export PATH=/opt/lotus/bin:/notes2/data:$PATH`

    (/notes2/data is the Notes data directory.)

16. Install the Domino data files and configure the Domino server ibiza/SEAT.

    Log in to ibiza as user notes1 and issue the command `notes`.

    Follow the steps in 3.1.6, "Setting Up the Domino Server" on page 34 to
    configure your Domino server according to your needs.

    Important

    Set the PATH field in Figure 4 on page 36 to `/notes1/data`. Remember
    that these servers are installed as partitioned Domino servers.

    ---

17. Install the Domino data files and configure the Domino server
    taurus/FORD.

    Log in to taurus as user notes2 and issue the command `notes`.

    Follow the steps in 3.1.6, "Setting Up the Domino Server" on page 34 to
    configure your Domino server according to your needs.

> **Important**
>
> Set the PATH field in Figure 4 on page 36 to `/notes2/data`. Remember that these servers are installed as partitioned Domino servers.

We assume that you now have your Domino servers up and running, configured as partitioned Domino servers, and each one starting up and listening on a different IP address. This IP address must be used for the service adapters of ibiza and taurus when you configure HACMP.

18. Once you finish the configuration of your Domino servers, start them to verify that they are working correctly. Check that you can start and stop them without problems.

19. Stop the Domino server Ibiza/SEAT in the system ibiza.

20. Stop the Domino server Taurus/FORD in the system taurus.

21. If you have the Notes workstation running, stop it in both systems.

22. Synchronize the volume group information of the shared volume groups in ibiza and taurus:

   - In system ibiza issue the command `umount /notes1`.

   - In system ibiza issue the command `varyoffvg notes1_vg`.

   - In system taurus issue the command `umount /notes2`.

   - In system taurus issue the command `varyoffvg notes2_vg`.

   - In system ibiza issue the command `importvg -y notes2_vg hdiskX`.

   - In system ibiza issue the command `chvg -a'n' notes2_vg`.

     (Where hdiskX is one disk from the notes2_vg taurus volume group.)

   - In system ibiza issue the command `varyoffvg notes2_vg`.

   - In system taurus issue the command `importvg -y notes1_vg hdiskY`.

   - (Where hdiskY is one disk from notes1_vg ibiza volume group.)

   - In system taurus issue the command `chvg -a'n' notes1_vg`.

   - In system taurus issue the command `varyoffvg notes1_vg`.

23. In both systems, check that:

   - The shared volume groups notes1_vg and notes2_vg have the field Activate volume group automatically at system restart set to `no`.

   - The shared file systems /notes1 and /notes2 have the field Mount automatically at system restart set to `no`.

- The journal file system log logical volume names for the shared volume groups notes1_vg and notes2_vg are different.
- The shared volume groups notes1_vg and notes2_vg are varied off in both systems.

24. Set up your TCP/IP configuration. This includes:

- Configure the file /etc/hosts in both systems. Even if you are using a DNS server for name resolution, you should have a local /etc/hosts file to avoid name resolution problems if your name server goes out of service. Here is an example of the entries in /etc/hosts in our systems ibiza and taurus:

```
192.168.110.23  ibiza_boot
192.168.110.13  ibiza
192.168.200.13  ibiza_sb

192.168.110.25  taurus_boot
192.168.110.15  taurus
192.168.200.15  taurus_sb
```

Figure 36. File /etc/hosts for the Taurus and Ibiza Systems

- Configure the file /.rhosts in both systems. Here is an example of the entries in /.rhosts in our systems ibiza and taurus:

```
taurus
taurus_sb
taurus_boot
ibiza
ibiza_sb
ibiza_boot
taurus_boot.msc.itso.ibm.com
taurus_sb.msc.itso.ibm.com
taurus.msc.itso.ibm.com
ibiza_boot.msc.itso.ibm.com
ibiza_sb.msc.itso.ibm.com
ibiza.msc.itso.ibm.com
```

Figure 37. File /.rhosts for the Taurus and Ibiza Systems

- Set up the standby adapters of ibiza and taurus. These network adapters must be initially configured with the standby IP address/mask. To check whether they are properly configured, issue the commands:

  rsh taurus_sb

  rsh ibiza_sb

If you did not login automatically to taurus_sb and ibiza_sb, check your configuration again.

- Set up the service/boot adapters of ibiza and taurus. These network adapters must be initially configured with the boot IP address/mask. Ping the boot IP addresses to check your configuration. To check whether it is properly configured, issue the commands:

  `rsh taurus_boot`

  `rsh ibiza_boot`

  If you did not login automatically to taurus_boot and ibiza_boot, check your configuration again.

25. Set up the HACMP configuration.

   In this step, we go through the HACMP for AIX SMIT panels that allow you to configure the HACMP part. You only need to do the whole configuration in one of the HACMP nodes. Once it is finished, you can synchronize with the rest of the HACMP nodes.

   1. Log into system ibiza as user root.

   2. Issue the command `smitty hacmp`. Select **Cluster Configuration** -> **Cluster Topology** -> **Configure Cluster** -> **Add a Cluster Definition**. We set the following values:

      - Cluster ID: `6` (it can be a number from 0 to 9999999999)
      - Cluster Name: `Cars`

   3. Issue the command `smitty hacmp`. Select **Cluster Configuration** -> **Cluster Topology** -> **Configure Nodes** -> **Add Cluster Nodes**. We set the following values:

      Node Names: `ibiza taurus`

   4. Issue the command `smitty hacmp`. Select **Cluster Configuration** -> **Cluster Topology** -> **Configure Adapters** -> **Add an Adapter**. We added the adapters ibiza, ibiza_boot, ibiza_sb, taurus, taurus_boot, and taurus_sb with the following values:

      [ibiza] (service adapter for node ibiza)

      - Adapter IP Label - `ibiza`
      - Network Type - `ether`
      - Network Name - `carether`
      - Network Attribute - `public`
      - Adapter Function - `service`
      - Adapter Identifier - `192.168.110.13`
      - Adapter Hardware - `Address MAC_Address_Of_Card`

- Node Name - `ibiza`

[ibiza_boot] (boot adapter for node ibiza)

- Adapter IP Label - `ibiza_boot`
- Network Type – `ether`
- Network Name – `carether`
- Network Attribute - `public`
- Adapter Function – `boot`
- Adapter Identifier – `192.168.110.23`
- Adapter Hardware - Address (empty)
- Node Name – `ibiza`[ibiza_sb] (standby adapter of node ibiza)
- Adapter IP Label -`ibiza_sb`
- Network Type -`ether`
- Network Name - `carether`
- Network Attribute - `public`
- Adapter Function – `standby`
- Adapter Identifier – `192.168.200.13`
- Adapter Hardware Address - (empty)
- Node Name - `ibiza`

[taurus] (service adapter for node taurus)

- Adapter IP Label - `taurus`
- Network Type – `ether`
- Network Name – `carether`
- Network Attribute - `public`
- Adapter Function – `service`
- Adapter Identifier – `192.168.110.15`
- Adapter Hardware – Address `MAC_Address_Of_Card`
- Node Name – `taurus`

[taurus_boot] (boot adapter for node taurus)

- Adapter IP Label - `taurus_boot`
- Network Type – `ether`
- Network Name – `carether`
- Network Attribute – `public`
- Adapter Function – `boot`
- Adapter Identifier – `192.168.110.25`
- Adapter Hardware Address - (empty)
- Node Name – `taurus`

[taurus_sb] (standby adapter of node taurus)

- Adapter IP Label - `taurus_sb`
- Network Type – `ether`
- Network Name – `carether`

- Network Attribute - `public`
- Adapter Function – `standby`
- Adapter Identifier – `192.168.200.15`
- Adapter Hardware Address - (empty)
- Node Name - `taurus`

5. Issue the command `smitty hacmp`. Select **Cluster Configuration** -> **Cluster Resources** -> **Define Resource Groups** -> **Add a Resource Group**. We created two resource groups and applied the following values:

   [ibiza]

   - Resource Group Name `ibiza`
   - Node Relationship `cascading`
   - Participating Node Names `ibiza taurus` (keep the order)

   [taurus]

   - Resource Group Name `taurus`
   - Node Relationship `cascading`
   - Participating Node Names `taurus ibiza` (keep the order)

6. Issue the command `smitty hacmp`. Select **Cluster Configuration** -> **Cluster Resources** -> **Define Application Servers**. We created two application servers and set the following values:

   [ibiza]

   - Server Name `ibiza_domino`
   - Start Script `/custom_scripts/start_ibiza_domino`
   - Stop Script `/custom_scripts/stop_ibiza_domino`

   [taurus]

   - Server Name `taurus_domino`
   - Start Script `/custom_scripts/start_taurus_domino`
   - Stop Script `/custom_scripts/stop_taurus_domino`

   We recommend that you create your own script directory for the scripts and utilities that you create. In our example, we used different scripts to start/stop the Domino servers in ibiza and taurus, but you can have the same start/stop script for all your Domino servers. However, in some situations you will need to start/stop your servers with different procedures.

7. Issue the command `smitty hacmp`. Select **Cluster Configuration - Cluster Resources - Change/Show Resources for a Resource Group**. We changed ibiza and taurus and set the following values:

   [ibiza]

- Service IP label `ibiza`
- Filesystems `/notes1`
- Application Servers `ibiza_domino`
- (the rest by default)

[taurus]

- Service IP label `taurus`
- Filesystems `/notes2`
- Application Servers `taurus_domino`
- (the rest by default)

8. Issue the command `smitty hacmp`. Select **Cluster Configuration** -> **Cluster Resources** -> **Change/Show Run Time Parameters**. We changed ibiza and taurus and we set the following values:

[ibiza]

- Debug Level `high`
- Host uses NIS or Name Server `true`
- Cluster Security Mode `Standard`

[taurus]

- Debug Level `high`
- Host uses NIS or Name Server `true`
- Cluster Security Mode `Standard`

9. Now our HACMP configuration is finished. To synchronize the HACMP resource definitions in taurus, from ibiza issue the command `smitty hacmp`. Select **Cluster Configuration** -> **Cluster Resources** -> **Synchronize Cluster Resources**. Run it with the default values. It must complete without errors. If you get an error, correct it and synchronize again.

10. To synchronize the HACMP cluster topology in taurus, from ibiza issue the command `smitty hacmp`. Select **Cluster Configuration** -> **Cluster Topology** -> **Synchronize Cluster Topology**. Run it with the default values. It must be complete without errors. If you get an error, correct it and synchronize again.

11. Now we can save our HACMP configuration for recovering in case of problems. Issue the command `smitty hacmp`. Select **Cluster Configuration** -> **Cluster Snapshots**.

26. Now the HACMP cluster is ready to run. You can start up the Cluster Services in both systems, try a Graceful Takeover for both systems and so forth, just to test that your configuration is correct.

3.5 Implementation of HACMP ES

HACMP ES is an enhancement of HACMP for AIX specially designed for the RS/6000 SP. It takes full advantage of the High Availability Infrastructure (HAI), which is part of the Parallel System Support Program (PSSP).

The implementation of an HACMP ES environment is very similar to the one described in 3.4, "Implementation of HACMP for AIX" on page 73. However, some differences make this topic necessary.

The way HACMP ES interacts with the High Availability Infrastructure is depicted in Figure 38 on page 85.
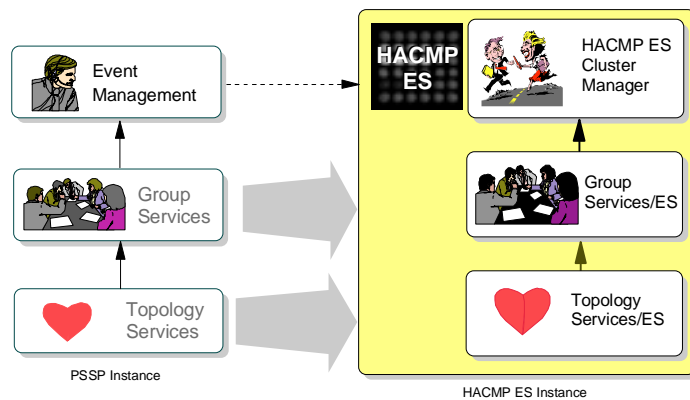
Figure 38. Relationship between HACMP ES and HAI Components

As you can see in Figure 38, HACMP ES uses special instances of Topology Services and Group Services daemons. In HACMP for AIX, these tasks are imbedded in the Cluster Manager daemon.

One of the main differences between the two HACMP implementations is the way they detect failures in the cluster. While HACMP for AIX uses an internal heartbeating mechanism, HACMP ES uses Group Services/ES, which provides faster detection for nodes going up or down, and also scales higher than the internal HACMP for AIX mechanism. Also, HACMP ES uses external agents, such as Event Management, for all other events. This last difference makes HACMP ES highly customizable, giving the possibility of using the more than 360 resource variables available in PSSP.

3.5.1 HACMP ES Specifics

In order to implement the environment described in Figure 35 on page 75, we can use two RS/6000 SP nodes sharing a common disk, which will contain the two file systems served by the two systems, as shown in Figure 39.
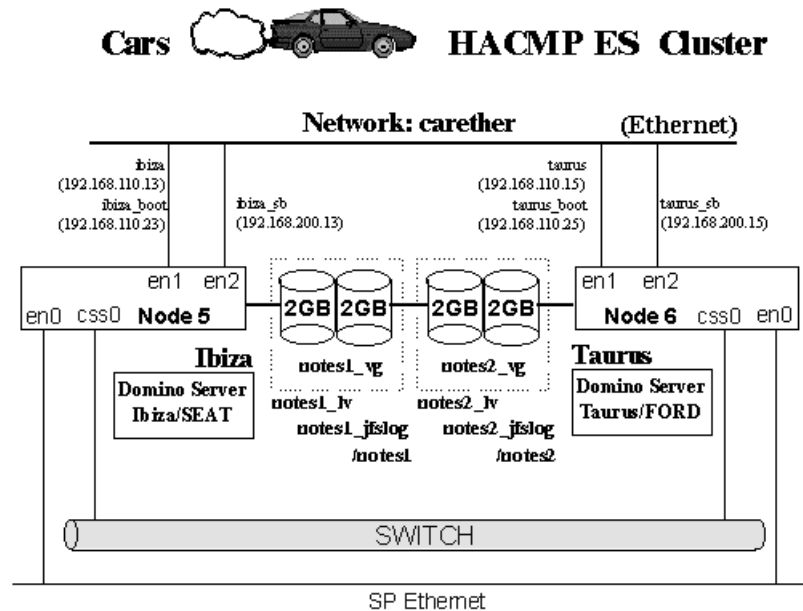
Figure 39. Mutual Takeover With HACMP ES

As you can see in Figure 39, the switch and the SP Ethernet are used to connect the two nodes. There is no need for an extra link between these nodes. The additional Ethernet adapter is a stand-by adapter used for taking over the service adapter from the other server.

The steps for implementing the HACMP ES cluster are exactly the same as those described in 3.4.1, "HACMP Configurations" on page 73. There is only one detail you should pay attention to: HACMP ES uses the information stored in the System Data Repository (SDR) for the Ethernet and switch adapters, so you should include all the adapters installed on every node that is part of the cluster, including the en1 adapter present as stand-by adapter, as part of the SDR information.

3.6 Keeping Your Domino Server Alive at the Application Level

As we explained in the introduction to topic 3.4, "Implementation of HACMP for AIX" on page 73, HACMP for AIX does not provide HA at the application level. Defining the start/stop scripts in the resource groups, as we showed in 3.4.2, "Example of HA Domino Configuration Using HACMP for AIX" on page 74, allows you to start/stop your applications automatically when a take over occurs, but it does not give you any more control over your applications.

The Domino clustering feature provides you with these HA capabilities:

- Cluster replication

  This enables you to maintain more than one copy of your critical databases.

- Cluster transparent failover

  Assuming cluster replication is running, if a client tries to open a database in a Domino server and the database is not available, the client will failover to a different replica of this database in the cluster.

These two HA features are useful but, depending on your needs, may not be enough for you. If the service that your Domino servers are providing to your users is really critical, you need to implement the highest possible HA configuration. You will need HA at the Domino server application level.

If your Domino server has a problem (the server process crashes, the Name and Address Book database gets corrupted, the mail routing does not work, and so forth), neither HACMP nor Domino clustering can do anything to correct it. You need additional HA software to fix these application failures.

To help meet your HA needs, we provide some examples and information on how to keep your Domino servers alive at the application level.

3.6.1 Simple Monitoring of the Response of Your Domino Server

This is an example of a simple way to add more HA in your Domino environment, monitoring the Domino server and recovering from server crashes. You can find the source of the C program and the shell script used for this example in Appendix A.1, "Scripts and Programs" on page 137.

In the following sections, we offer information about some more complex HA software solutions developed by software vendors, and some basic ideas from which you can develop your own solutions.

The idea is to start to add more capabilities to your HA Domino environment. The example of a shell script that will always be running in your Domino environment monitoring your Domino servers. The script gets a list of Domino servers that must be checked and, for each server, opens a Notes connection and reads a specific database of this server. If the access is successful, it means that the Domino server is up and running; otherwise, we consider that the server is down.

Considerations about the C program and the shell script:

- It is better to create a specific database in each Domino server and use it for the checking. For performance reasons, you should avoid the NAB and heavily used common databases.

- The name of the servers you put in the servers file can be the TCP/IP name of the Domino server or the Notes name of the server. For example, it can be server09 or server09/ITSO.

- We left the recovery actions part of the script intentionally blank. Here you can add your own restart_server scripts or any additional recovery procedure you need for your environment. Try not to make it complex.

- This example can be used in any situation, but it will be more efficient if used in conjunction with HACMP. For example, if you are using HACMP for AIX, you can run the shell script in a root environment (the C program must be run in a Notes environment) and force a graceful takeover in the system that runs the failed Domino server. It can be useful if your recovery actions cannot restart the Domino server.

As you can see, this is the basic and most simple way to start to monitor/control/recover your Domino servers.

This method of providing HA at the Domino server application level requires a programming and customizing effort if you want to implement a full HA Domino environment. In the following sections we provide information about some of the products available to achieve your HA Domino environment.

3.6.2 The Notes Recovery Kit (NRK)

The Notes Recovery Kit is a powerful solution, developed by IBM, to keep your Domino servers up and running. NRK maximizes the availability of Lotus Notes applications on standalone AIX and SP systems.

NRK is one of the most complete software package available to achieve the highest HA for your Domino environment.

> **Important**
>
> NRK is available from IBM only through a PRPQ (request for quotation) and it is not supported by Lotus.

The features of NRK are:

- It is built on top of the High-Availability Application Services (HAAS).

  It works on top of HACMP for AIX and subscribes to the Parallel System Support Programs (PSSP) Group Services in the RS/6000 SP.

  Each Domino server operates under its own HAAS service manager.

- In standalone configurations, the NRK provides only local recovery of Notes servers. Clustered configurations add support for failover of Notes servers between cluster hosts.

  In standalone mode, the NRK can support multi-server configurations, including Notes partitioning.

- It works with a Notes-specific recovery model.

  Multiple Domino components are modeled in the NRK, such as:

  - Server availability/responsiveness
  - Mail router
  - NAB management

- Multiple server instances are supported (partitioning), such as

  - Notes partitioning
  - A virtual local operator console for each server instance
  - Independent management of server instances

- It provides deep monitoring of the Domino servers (Notes Health Monitoring):

  - Connectivity
  - Functionality
  - Responsiveness
  - Local console responsiveness

The Notes Health Monitor (NHM) and the Notes Console Interface Monitor (NCIM) work together to automate the typical manual detection used in most Lotus Notes environments, in the following ways:

1. A client attempts to connect to the Lotus Notes server.

2. If the client encounters a failure, it notifies an administrator.

3. The administrator checks the Lotus Notes server console.

4. If the Lotus Notes server is unresponsive, the administrator restarts the server.

The Notes Health Monitor (NHM) helps you make sure Lotus Notes is running well by periodically probing specified Lotus Notes servers to ensure they are responsive to client requests. A probe connects with a Lotus Notes server socket port and reads a Lotus Notes database residing on the Lotus Notes server.

- Notes server restart feature, which offers the following:

  - Is triggered by NRK recovery model actions.
  - Resets/cleans application environments before restart.
  - Clients do not lose login.
  - Resubmits failed transactions.
  - Does not interfere with Notes 4.5x self-recovery features.

A Lotus Notes server may stop responding to client requests, or the Lotus Notes console may stop responding to commands. If this happens, restarting the server often corrects the problem that caused server unresponsiveness. Another type of problem is when the server process has failed. A quick restart is the objective.

The goals of detecting server unresponsiveness are:

  - Determining whether the Lotus Notes server is temporarily or permanently unresponsive within the time limit dynamically set at runtime.
  - Imposing minimal overhead on the system.

The amount of time you give the NRK to detect Lotus Notes server unresponsiveness depends on many system factors, including:

  - Notes server configuration
  - The performance characteristics of the hardware and operating system
  - The typical user load on the system
  - The tasks that are currently executing

- Name and Address Book (NAB) management.

  - Useful in environments with very large NABs (IBM's NAB is > 1GB).
  - Maintains a backup of the NAB for each server, and auto-replaces it in case of failure.
  - Provides background NAB replication capability. This is useful for reducing server loads. It is applicable in one-way NAB replication environments.

The NRK and HAAS restart Lotus Notes server applications after execution stops due to a problem. Sometimes, when a Lotus Notes server application stops executing under these circumstances, its NAB becomes

corrupted. If the Lotus Notes server application uses the corrupted NAB after restart, the corrupted NAB may cause further problems that stop execution.

You can specify the interval at which the NRK should back up the NAB for a particular Lotus Notes server application when you configure the associated Lotus Notes service.

The backup copy of the NAB resides on the same AIX host as the corresponding Lotus Notes server application, which saves time when the NRK must replace a corrupted NAB with a backup copy.

If the Lotus Notes server application fails too many times within the specified time limit for the associated Lotus Notes service, the NRK replaces the existing NAB of the Lotus Notes server application with the backup NAB.

In addition to replacing NABs that have already been corrupted, the NAB backup facility can be configured to periodically compact the NAB with the customizable script:

/usr/ha/class/NABMgr/bin/replicatenab

This compaction helps to prevent the NAB from being corrupted due to exceeding the Lotus Notes database file size limits or file system free space limits. Large NABs, in particular, benefit from being compacted.

- Lotus Notes Mail Router failure.

    When a mail router detects corruption in the mail.box database, it fails and shuts down cleanly. The NRK then:

    - Detects that the mail router process terminated.
    - Runs a consistency check on the mail.box file.
    - Instructs the Lotus Notes Server to restart the mail router.

    If the mail router fails more than once within a previously specified length of time, the NRK concludes that the mail.box corruption is not correctable using fixup and that intervention is required by an administrator. The NRK renames the mail.box file to reflect when the failure occurred. For example, if the failure occurred on August 20, 1997 at 2:31 p.m. (or 1431 in 24-hour time format), the new filename is:

    mail.box.0820_143100

    The NRK then sends a notification to the administrator using a pager or e-mail.

    When the mail router restarts, the router detects that there is no mail.box file and creates a new one. An administrator must try to recover

documents from the old, renamed mail.box file by manually copying and pasting them.

- Clustering support with failover/takeover/migration capabilities.

  Support for highly available environments where clients are dinamically routed to the next available copy of the database. NRK is:

  - Complementary to Notes clustering.
  - Fully integrated with HACMP for AIX.

    A Lotus Notes server configured using the NRK under the HACMP offers all the benefits of an HACMP environment plus the benefits of Lotus Notes server's recovery capabilities with the NRK.

    Important

    You can configure NRK to operate with HACMP rotating or cascading mode. HACMP with NRK is not supported under HACMP concurrent mode or in a Network Information System (NIS)

  - Fully integrated with PSSP Group Services (PGS) for RS/6000 SP systems.

    A Lotus Notes service configured to run in clustered mode using PSSP Group Services adds the following to a standalone configuration:

    - If an AIX host fails, each application running on that host is taken over by a host previously specified as the standby AIX host for that application. The standby AIX host acquires the resources to run the application and then runs it.
    - You can manually migrate an application from one host to another. You can move each application independently of others that are running on a particular host.

    In NRK (on top of Group Services), a Notes group is the set of nodes on which a particular application can run. A Notes group has the following characteristics:

    - It is a subset of a PGS cluster.
    - It is a dynamic entity; as hosts come up and go down (either by boot up, shutdown, failures, and so on), the group grows and shrinks.
    - HAAS and the NRK support a maximum of 2 nodes per group in a Group Services configuration. Takeover and migration operations for an application occur between pairs of nodes.
    - Each application that may have a takeover or migration performed on it must have its own Notes group defined. (When you define

Notes groups during the configuration of Lotus Notes services, these groups are used for the PGS configuration.)

- A host can be a member of multiple groups. For each group to which the host belongs, it can be paired with any of the other hosts in the PGS cluster. Consequently, when multiple applications are running on a host and the host fails, each of the applications can move to a different host.

Regarding to arbitration; at any time, an application runs only on one host in the Notes group. Arbitration is the process of determining the host on which the application runs. HAAS arbitration is built on the infrastructure provided by PGS. It resolves issues such as the following:

- Which host the application should run on if two hosts come up simultaneously
- Whether another host in the PGS group is already running the application when a host comes up
- Which host to select to run the application when a host fails

- Tivoli integration (future).

This software is an IBM PRPQ. Contact your IBM representative for more information.

3.6.3 TME 10 Module for Domino/Notes

If you are a Tivoli customer, you can benefit from the TME 10 Module for Domino/Notes.

The TME 10 Module for Domino/Notes allows for the management of Domino/Notes environments by providing the following functions:

- Security enhancements
- Task administration
- Availability management
- Automated Notes client installation
- User management provided with the Admin Extension for TME 10 Module for Domino/Notes

Since the TME 10 Module for Domino/Notes utilizes the TME 10 Framework and TME 10's other core applications, the functionality the module provides is fully customizable and extendable.

Because we are focussing on providing HA for the Domino environment, we only deal with information related to availability management.

Managing the availability of a Notes environment with the TME 10 Module for Domino/Notes includes the following features:

- TME 10 Event Adapter for Alert Monitoring

  The TME 10 Module for Domino/Notes provides a Notes add-in task in the form of an event adapter that monitors Notes alerts and translates them into Tivoli/Enterprise Console (T/EC) events.

- TME 10 Notes Monitor Collections

  The TME 10 Module for Domino/Notes supplies two customized monitor collections, Central and Remote. These collections check such things as:

  - Server availability
  - Notes statistics
  - Whether daemons, replicators, routers, and add-in tasks are running

  Proactive monitoring and automatic creation of different severity levels for T/EC events is enabled through configurable thresholds and responses.

- Event Filtering

  The TME 10 Module for Domino/Notes uses a configuration file to allow the administrator to filter selected events to prevent them from being forwarded to the Event Server Rulebase for processing.

- Tivoli Event Server Rulebase

  The Tivoli Event Server Rulebase provides intelligent event processing and correlation. The TME 10 Module for Domino/Notes provides two rule sets to simplify and enhance Notes management:

  1. The T/EC adapter rule set captures and processes Notes events from the TME 10 Event Adapter and forwards them to the TME 10 Enterprise Console.
  2. The TME 10 Distributed Monitoring rule set receives and processes information from the Notes Central and Remote monitoring collections and forwards them to the TME 10 Enterprise Console.

  These rule sets enable such features as automatic closing of harmless events, event correlation, and multi-region support.

  For more information on the TME 10 Module for Domino/Notes, go to:

  `http://www.tivoli.com`

3.6.4 The PATROL Knowledge Module for Lotus Domino/Notes

If you are already a PATROL user, you can benefit from the PATROL knowledge module for Lotus Domino/Notes.

The PATROL Knowledge Module for Lotus Domino/Notes improves the availability and performance of Lotus Notes and Domino.

This software:

- Monitors Domino servers for continuous availability.
- Ensures that mail/replication processes are working at best performance.
- Tracks the size and performance of Lotus Notes databases.
- Makes Domino administration and maintenance easier.

Because our intent is to provide an HA Domino environment, we only discuss the HA capabilities of this product.

The features for monitoring and managing Domino servers include:

- Monitoring Domino servers for availability and alerts the administrator of problem and potential problem areas.
- Gathering statistics that let you monitor neighboring servers and know that necessary connections are working.
- Measuring response time in number of seconds to determine if there are performance problems.
- Automatically implementing PATROL recovery actions when problems occur (for example, automatically restarting the server), or sending e-mail to notify the administrator if manual attention is required.

The features for monitoring mail and replication tasks include:

- Ensuring the router is working, tracking the number of messages being routed and the number of messages that were not transferred.
- Monitoring the mail.box database for statistics such as large numbers of undelivered mail messages and messages waiting too long in the queue.
- Monitoring the size of users' mail databases for those that are larger than a user-defined maximum.

For more information about this PATROL knowledge module, go to:

`http://www.patrol.com/products/pat/km/plnds.html`

Chapter 4. Performance Tests

In this chapter we explain the performance results that we obtained from some specific performance tests done in our lab environment.

The purpose of these performance tests was to compare standalone Domino installations versus partitioned Domino installations in SMP systems, and determine the performance impact when going from a standalone Domino installation to a Domino cluster. These performance tests did not provide Domino server sizing information.

The tool used for the tests was NotesBench, and the NotesBench workloads that we used were the Mail and Shared Database workload (maildb) and the Cluster Impact Topology workload (cluimpact).

Important

To be able to understand our explanations and the test results, you should have a fair knowledge of NotesBench.

- We used the maildb workload to compare the performance between a standalone Domino server and a set of partitioned Domino servers using the same SMP system. This workload models a server for active users who are only performing mail and simple shared database operations. With this test, we tried to demonstrate that you can benefit from partitioned Domino configurations running on SMP systems. The load given to the standalone Domino server was exactly the same as that given to the partitioned servers. The reason for using this test is that it simulates a common workload.

  Each simulated user of the workload runs the Mail Test script, which models an active user on a client reading and sending mail. It contains an average of 15 minutes of waiting, which means that an average user executes this script no more frequently than four times per hour. For each iteration of the script, there are five documents read, two documents updated, two documents deleted, one view scrolling operation, one database opened and closed, one view opened and closed, and some miscellaneous operations. One message is sent to the recipients approximately every 90 minutes.

  None of the messages sent by each driver user are delivered to any mail database on the system under test. The updates done in the Mail Test are intended to model reading and submitting messages, but not message delivery or reading delivered messages.

97

The measurements obtained by this test are:

- Throughput of completed Notes operations
- Average response time at maximum capacity
- Maximum mail users supported

The resultant capacity metric for a mail-only server is the maximum number of users that can be supported before the average user response time becomes unacceptable, but we remind you that we are only comparing performance results.

- We used the cluimpact workload to show what the performance impact is when you go from a standalone Domino configuration to a Domino cluster. This workload is a three-phase workload that models a server that initially is a standalone mail and shared discussion server, then becomes part of a cluster, and finally exercises server failover. In each phase, the maildb workload is used. Each phase generates independent performance information. The reason for using this test is that it tests the complete path from standalone server to cluster, including failover.

Disclaimer

In regard to NotesBench for Lotus Notes R4.6.1:

We chose NotesBench for Lotus Notes R4.6.1 because it is the standard tool for Lotus Notes benchmarking.

The people who installed, configured and ran NotesBench were not certified by the NotesBench consortium.

The performance tests we ran are not intended to provide server sizing information.

The setup of the benchmarks was done by strictly following the guidelines given in the NotesBench documentation.

For more information about the NotesBench software (different tests, different workloads, and so forth) go to:

http://www.notesbench.org

4.1 Cluster Tests

Using the raw data from several runs on the systems, the following is a summary of the results from running Probe, Groupware_b, Cluster Mail and Cluster Impact NotesBench tests on the Poughkeepsie SP systems.

The objective of the tests was to evaluate the performance of one Notes server system vs. three Notes servers in a cluster.   Testing was completed using two SP frames, with AIX 4.3 and Notes 4.6.1. These systems had a token-ring network for client traffic. They also had a high-speed switch in some of the tests, used for server-to-server communications.

Cluster Overhead - Some overhead is associated with enabling the cluster tasks. Tests showed approximate response times to be 10 to 18% slower when the cluster tasks were enabled, using a light load on the system.

Some network traffic occurs in support of cluster communication, but this traffic has almost no impact on overall performance. The network cost was usually 2 transactions per minute per server in the cluster.

Multiple System Advantage - There are performance advantages to having multiple systems when the single system is nearing capacity. In the single server test, the CPU utilization was 90% or above when running the heavy load test, and the system showed a slight amount of paging activity. When the same work was executed against three servers in a cluster, the servers' CPU utilization went to 50, 55 and 60% on those three servers. Client-to-Server network traffic remained the same, and moving from one network connection to three network connections spread out the work and reduced network card contention. Server-to-Server event replication increased the network load. In the heavy load test, the server-to-server network traffic was approximately the same as the client-to-server traffic. When running on one network, the network traffic was slightly more than double due to event replication.

Response Time - Response time improved when three clustered servers were used instead of one server, in heavy load testing, only when the workload on the original server had that system near capacity or overcommitted (CPU, memory or disk). The advantage of a cluster is in data availability through replication, and improved performance for clients when work can be balanced across multiple systems from an overcommitted system.

Single Server Performance - Response time degraded severely when the server was overcommitted (CPU, memory or disk) with a heavy workload.

Subsecond response times became several seconds, with as much as 14 and 18 seconds observed, when the workload on the server was heavy. When the single server was not overcommitted prior to being added to the cluster, response time degraded by up to 20% when that server was added to the cluster.

Clustered Server Performance - Response times degraded when the server was either replicating or indexing. In this experiment, there was only one replicator task even when there were three servers. It is possible that having multiple instances of the replicator task can help, but this theory was not tested.

Failover - When failover occurred, response times degraded. It was also noted that there were a significant number of failed client requests when the test was running with failover occurring. This anomaly needs further study, as there appeared to be no reason why the requests could not have been serviced by the remaining two servers.

The response time degradation was approximately 1/4 second (14%) per client request. This degradation was relative to having all three servers available in the cluster. The response time of the three server cluster with one failed server running the heavy load test was still faster than the response time of the single non-clustered server.

Response Time Data - The average response times are shown in Table 10 and Table 11.

Table 10. Average Response Times - No Cluster

Single Server, no cluster	Time
Light load	145 milliseconds
Heavy load	14.1 seconds

Table 11. Average Response Times - Cluster

Three-Server Cluster	Time
Heavy load	181 milliseconds
Heavy load (w/replication task active)	250 milliseconds
Heavy load (w/failover)	207 milliseconds

4.2 Performance In Partitioned Domino Servers

The goal of testing performance in partitioned servers is to find out if a given hardware delivers more or less user performance when installed with Domino Clustering. We used a fixed load of 900 users on a given server. In the first test, all 900 users were served by the same server. In the second test, the 900 users were divided into 450 users on each of the two partitions running on the same server. In third test, the 900 users were divided into four server partitions. We were using five clients to generate the load.

The test did not push the system to its limits, but we tried to measure if a given system (4-way SP) with a given load (900 mail users) has a higher or lower NotesMark number with two or four Domino partitions than a server with just one partition.

The performance numbers we saw are only relative numbers, since the test was not benchmark. We did not have access to a hardware configuration with sufficient disks and LAN bandwidth for performance measuring or benchmarking. We had access to high-performance CPUs, one SP node with four 233MHz 604 CPUs and 1GB of memory. However, the system under test was not balanced; it had much more CPU capacity than needed for the I/O configuration, so these numbers can only be used to tell if, for this configuration, the performance will increase or decrease with more partitions.

Our test environment consisted of the server, five test clients, and a test parent. The clients and the parent were placed on another LAN segment, and test data had to cross four routers before reaching its final destination. These LAN conditions may have added some distortion to the response time numbers.

The reason for choosing the Mail Test was that it is a very typical workload for a customer. The reason for only simulating 900 users was simply our limitation in disk space. Each user requires a mail file (user.nsf), and we only had only disk space for 900 mailboxes.

4.2.1 Results of the Benchmarks.

The following results were recorded:

 • SP system with one partition

```
Test Run: Users = 721 NotesMark = 1859 Response Time = 359 msec
```

 • SP system with two partitions

```
Test Run: Users = 892 NotesMark = 2217 Response Time = 339 msec
```

• SP system with four partitions

```
Test Run: Users=899  NotesMark=2215 Response Time=323 msec
```

Conclusion of this set of benchmarks - From these results we find that there is a performance improvement by going to more partitions.

We concluded that the system did more work when installed with partitioned servers and therefore delivered better user performance. From AIX system monitoring (see typical vmstat output for each test in the following screens), we can see that the system load is I/O-bound, and the disk capacity seems to be the limit.

Test with one Domino partition in the server:

```
Silver one partition
kthr     memory               page               faults      cpu
----- ----------- ------------------------ ------------ -----------
 r  b   avm   fre  re  pi  po  fr   sr  cy  in   sy  cs us sy id wa
 0  0 77681   127   0   0   0   4    7   0 481  862 783  6  2 70 22
 1  2 77831   129   0   0   0  24   44   0 3345 9314 5722 14  5 37 43
 1  2 77885   129   0   0   0  24   44   0 3319 9048 5711 14  5 58 23
 1  6 78009   125   0   0   0  46   78   0 3407 9673 5733 16  7  6 72
```

Test with two Domino partitions in the server:

```
Silver two partitions
kthr     memory              page             faults      cpu
----- ----------- ------------------------ ------------ -----------
 r  b   avm   fre  re  pi  po  fr   sr  cy  in   sy  cs us sy id wa
 0  3 58748   123   0   0   0   2    8   0 1479 3713 2058  8  3 20 70
 0  2 58749   125   0   0   0   3    5   0 1419 3585 2010  8  2 62 28
 0  3 58749   128   0   0   0  10   24   0 1472 4366 2089 10  3 25 62
 0  3 58750   122   0   0   0   3    5   0 1471 3848 2067  9  3 29 60
 0  2 58751   125   0   0   0  11   20   0 1438 3967 2039  8  3 43 46
```

Test with four Domino partitions in the server:

```
Silver four partitions
kthr      memory              page                faults      cpu
----- ----------- ------------------------ ------------ -----------
0  3 58752   124   0   0   0    5    10   0 1458 4438 2085  9   3 26 62
0  2 58752   134   0   0   0    9    19   0 1437 3589 2036  7   2 33 57
2  9 81911   134   0   0   0   41    78   0 2342 16547 6457 44 11  0 46
2 12 82015   139   0   0   0   46   101   0 2277 12552 4267 31  7  0 62
0  2 58754   128   0   0   0    5    13   0 1389 2883 1970  6   2 75 17
0  3 58754   123   0   0   0    4     7   0 1447 3415 2051  8   3 32 58
```

Chapter 5. Maintaining a Domino Cluster

You can use the cluster capabilities of failover and workload balancing together to achieve greater database availability. There are different strategies for optimizing database and server access depending on the type of database and the use of it.

- Workload balancing

  Workload balancing is used to distribute the workload in the cluster. The workload is balanced by use of the parameters Server_Availability_Threshold and Server_MaxUsers in the notes.ini file.

- High failover

  User mail databases should be highly available. You can make a mail database more available by creating a replica on another server. In this way, users fail over to the replica if the mail file is unavailable. Because only one person uses a personal mail database, you do not need to workload balance the database servers. Clustering your mail databases will increase the load on the servers.

- Active workload balancing

  If you have a large number of user mail files on your system, you should consider distributing those files equally across the cluster so that no single server has a significantly larger number of mail files than any other server. This will spread the workload among the servers.

- High failover and active workload balancing

  When you have a database that users access continuously, such as a Web database or a special discussion database, not only do you need high availability, but you also need to distribute within the cluster the heavy workload created by users accessing the database.

- How do I set up failover and workload balancing to work together?

  As the cluster administrator, you should closely monitor your system for usage patterns during peak periods and adjust cluster resources accordingly. Provide failover for databases by creating replicas on two or more servers. You can balance server workloads by adjusting the Server_Availability_Threshold and Server_MaxUsers limits of each server. The cluster manager then monitors each server in the cluster, triggers failover, and redirects requests accordingly. The cluster manager also monitors the number of active users and indicates to other cluster servers when a server is in the MAXUSERS state and is not accepting new user requests.

As a system administrator you would also sometimes need to exclude users from accessing the server. The way to do this is to change the value of Server_restricted to 1 or 2.

5.1 Managing Servers in a Cluster

This section describes procedures and actions in the cluster maintenance process.

5.1.1 Removing Servers from a Cluster

You must have Author access or above in both the Public Address Book and the Administration Requests database, and Delete Document rights in the Address Book. If possible, use the administration server when removing a server from a cluster.

When you ask to remove a server from a cluster, the Administration Process, at the administration server, makes the change to the Server document in the administration server Address Book.

Domino submits this request to the Administration Requests database on the server from which you initiated the Remove from the cluster request. A separate request is submitted for each server that you remove from the cluster.

If you make this request on the administration server, the Administration Process immediately acts on the request. If you make the request on another server in the cluster, the admin4.nsf database must first replicate with the administration server admin4.nsf database before the Administration Process can run.

After the ADMINP task makes the changes to the removed-server Server document in the administration server Address Book, the changes must be replicated to the removed server. When the removed server receives the changes in the Address book, the cluster processes on the server exit.

5.1.2 Activating Servers in a Cluster

When a server has been stopped, and then reactivated, the server tries to synchronize all cluster databases, and you must be prepared for the load from this update in the system. Therefore, bring back servers during evenings or other non-busy times of day. After bringing a sever back to service, you should also be aware that cluster replication events are kept in memory on other servers. If a server has been out of service for a long time, some events

may be deleted from memory on other servers in the cluster. Therefore, if a server is put back in service, it might not be fully synchronized by the cluster replication process, and you must also issue a scheduled replication.

5.2 Managing Databases in a Cluster

This section describes how to manage databases, including adding databases and removing databases from cluster service. It also describes why we should ask the administration process to delete databases.

Before you begin to create databases and replicas in a cluster, consider how frequently users access a database and evaluate their need for data availability. For heavily used databases, you may want to create more than one replica and locate these on your most reliable servers.

In general, the more replicas of a database, the more accessible the data. Creating too many replicas, however, can add unnecessary maintenance overhead and therefore impact performance. As you plan your cluster implementation, you must balance user requirements for data availability and server workload.

In situations where users require the constant availability of a specific database, consider placing replicas on every member in the cluster, provided you have adequate disk space and resources to do this. If you are a public service provider, this configuration provides the highest possible redundancy of data.

Not all databases require multiple replicas in a cluster. A server log file, for example, should not have a replica on another server, because if the server fails, it does not need the log file.

To view existing databases and replicas in your cluster, open the Cluster Database Directory. Domino maintains a profile of all databases and replicas within a cluster in this directory database.

There are five database settings that let you manage databases in a cluster:

- Cluster Replication enabled
- Cluster Replication disabled
- Out of service
- In service
- Pending delete

The system administrator should control cluster replication by setting the value in the cluster directory data base. By disabling cluster replication for databases where it is not needed, you minimize the impact on the cluster performance when changes are made to these databases. Also, it may have some positive effect on the cluster if you shorten the list of databases the server has to monitor. To change the Cluster Replication setting, open the cldbdir.nsf database, find the document for your database, and set the value to enable/disable.

If you want to prevent users from accessing a database temporarily, you can mark it out of service. This is done from the menu by choosing **File**-> **Tools** -> **Server Administration** -> **Database Tools.**

When you mark a database out of service, users cannot open the database. New open database requests fail over to a replica. If there is no available replica, access to the database is denied and the message "`Access to the database has been restricted by the administrator`" appears.

However, Domino maintains all existing connections to the database and allows replication updates from other replicas to continue to occur. If users close an out-of-service database and then try to reopen it, the request is sent to another replica on an available server. This means that the database is gradually brought to an out-of-service state without disruption to users who are actively using it.

If you want to make all the databases on a server out of service, you can, as an alternative, set config server_restricted=1 at the server console. This setting causes the server to be restricted. A restricted server cannot accept new database open requests. If possible, requests are failed over to other servers in the cluster. However, existing connections to a database will not be terminated.

If you have marked a database out of service and now want to restore access to the database, you must mark the database in service. When you put an out-of-service database back into service, it becomes fully available to users.

If a database is obsolete and you need to delete it, you must do this with a minimal disruption to users. You can set a database to be deleted only after every active user has finished using the database. When you mark a database for deletion, no new database open requests are accepted because the database is immediately marked out of service. After all users have terminated their connections to the database, changes are pushed to another

replica and the database is deleted by the server. You should mark an obsolete database for deletion if you plan to remove it, or if you are copying the database from one system to another and want to delete the database from the original server.

You control the database attributes In Service, Out of Service and Pending delete by choosing **File** -> **Tools** -> **Server Administration** -> **Database tools**. Select the Server, Database, Cluster Management and mark a database out of service, in service or pending delete.

5.3 Logging in a Cluster

This section describes the logging performed by the cluster manager and cluster replication tasks.

For viewing the cluster manager events or the cluster replicator events, open the log database.

For analyzing the events logged in the log database, use the log analysis tool.

5.3.1 Log Database

The cluster manager events and the cluster replicator events are logged in the log.nsf database.

5.3.1.1 Cluster Manager Events

The cluster manager task logs failover and workload balancing information as entries in the Miscellaneous Events of the log database.

Figure 40 shows workload balancing entries in the log file.

Figure 40. Cluster Manager Events (Partial View)

5.3.1.2 Cluster Replicator Events

The cluster replicator task logs information about replications that completed successfully, and also about replications that failed to complete and are waiting for retry.

This information is logged as Replication Events in the log database. To find all the events generated by the cluster replicator, search for the keywords CLUSTER REPLICATOR in the field Initiated By of the log database.

Figure 41 shows the cluster replicator events in the log file.

With Server	Date	Starting Time	Ending Time
▼ servnt2			
	▶ 16/04		
	▶ 17/04		
	▼ 18/04		
		10:56	10:56
		11:17	11:17
		11:37	11:37
		11.56	11.56

Notes Log (Server01/PARTITIONED)
- Database
- Mail Routing Events
- Miscellaneous Events
- NNTP Events
- Phone Calls
- Replication Events

Server01/PARTITIONED
18/04 10:56 – 18/04 10:56

Remote Server: ServNT2/PARTITIONED
Initiated By: CLUSTER REPLICATOR ON SERVER01/PARTITIONED
Elapsed Time: 0 minutes

Events
Unable to replicate from cldbdir.nsf to ServNT2/PARTITIONED cldbdir.nsf: Server not responding
Unable to replicate from admin4.nsf to ServNT2/PARTITIONED admin4.nsf: Server not responding

Database	Access	Added	Deleted
ServNT2 statrep.nsf	Manager	124	70

Figure 41. Cluster Replicator Events (Partial View)

The cluster replicator maintains the replication information in memory for one hour before logging it. For this reason, only replications that have failed to complete their retries during this period are logged.

To force a manual log at any time, use the following command in the server console:

```
tell clrepl log
```

5.3.2 Log Analysis Tool

Use the log analysis tool available in the administrative panel to search the cluster manager or cluster replicator entries in the log database, as follows:

1. From the Notes client open the server administration panel, choose **File -> Tools** -> **Server Administration**.

2. Select the server whose log file you want to search.

3. Click the **Servers** icon and choose **Log Analysis**.

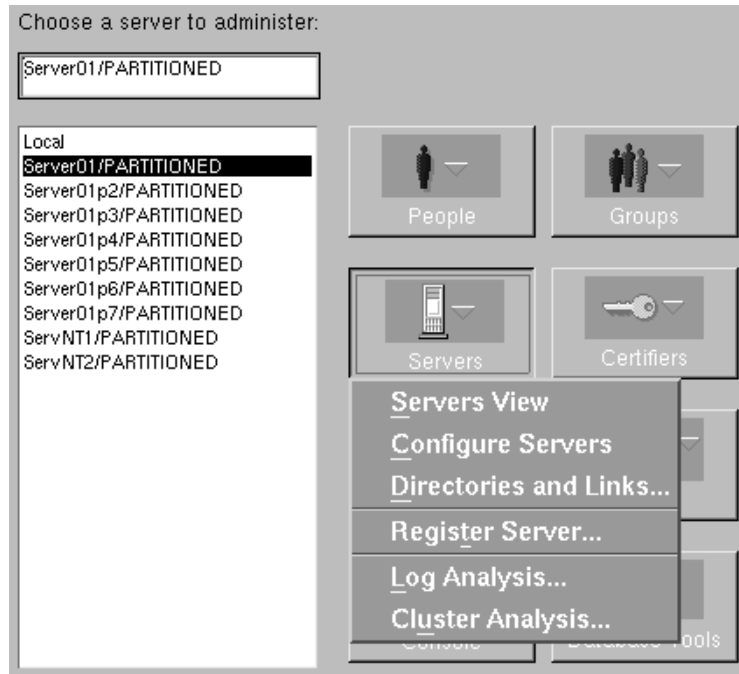   Figure 42 shows the selection of **Log Analysis** for Server01/PARTITIONED.
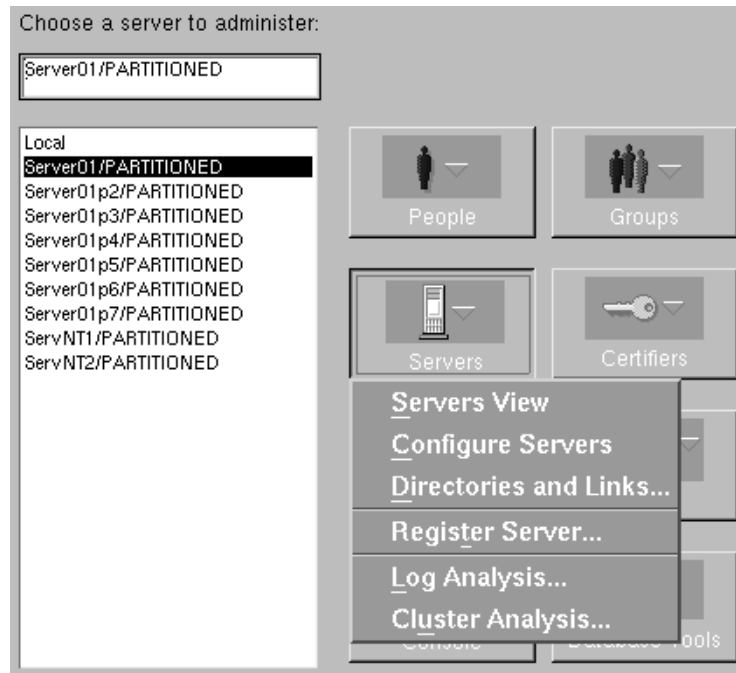
Figure 42. Log Analysis - Selecting Server from the Administration Panel

4. Click **Results Database**, select the server to store the database, and specify a title and a file name for it.

5. If a Results Database already exists, select **Overwrite this database** to remove any previous entries or select **Append to this database** to add the new entries to the existing ones. Click **OK**.

6. Enter the number of days to search.

7. Type individual words for which you want to search, separated by commas.

   Figure 43 on page 113 shows the selections for a local Log Analysis database with no search criteria for a period of one day.
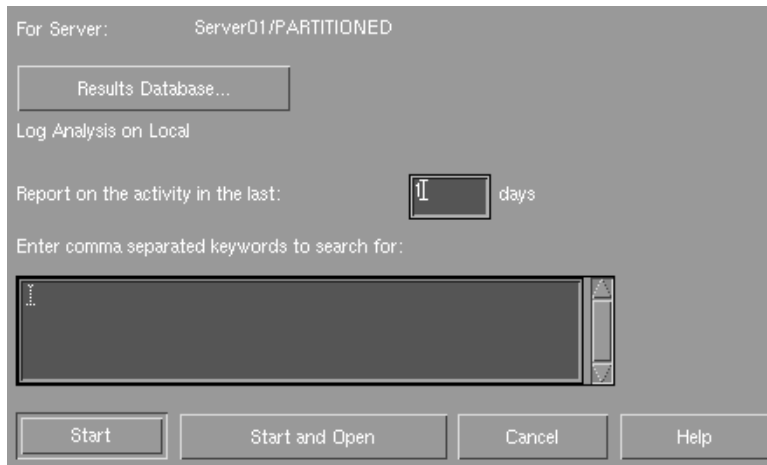
Figure 43. Log Analysis - Defining Result Database(Partial View)

8. Click **Start** to search the log file; or click **Start and Open** to search the log file and then immediately open the Results Database to view the results.

Figure 44 on page 113 shows the results database generated.

Figure 44. Log Analysis - Log Analysis Database (Partial View)

5.4 Statistics in a Cluster

This section describes how to check the cluster configuration and the cluster statistics provided by the cluster manager and cluster replicator tasks.

For checking the cluster configuration, you have the cluster analysis tool.

For viewing statistics provided by the cluster manager and the cluster replicator tasks, you have the following choices:

- Statistics commands at the server console
- Statistics reports in the statistics database

5.4.1 Cluster Analysis Tool

The cluster analysis tool is very helpful for checking the cluster settings after installation. The cluster analysis tool checks for correct server processes running, a consistent ACL list, replica copies, replication and so forth.

Doing a cluster analysis can be done in a few minutes, and will give you valuable information that can save you time.

To set up for the cluster analysis, do the following:

1. From the Notes client, open the server administration panel, choose **File -> Tools -> Server Administration**.
2. Select or enter the server name whose analysis you want to run.
3. Click the **Servers** icon and choose **Cluster Analysis**.

   Figure 45 on page 115 shows the selection of cluster analysis for Server01/PARTITIONED.

Figure 45. Cluster Analysis - Selecting Server From the Administration Panel

4. Click **Results Database** and specify the database to receive the results of the analysis test.

5. If a Results Database already exists, select **Overwrite this databas**e to remove any previous entries or select **Append to this database** to add the new entries to the existing ones. Click **OK.**

6. Select the types of reports you want generated: *Server*, *Database*, or *Server Web Navigator*. The *Server* and *Server Web Navigator* reports are defaults. If you select *Database*, you can also select *Database details*.

   Figure 46 on page 116 shows the selection for a local Cluster Analysis database with report details for Server only.
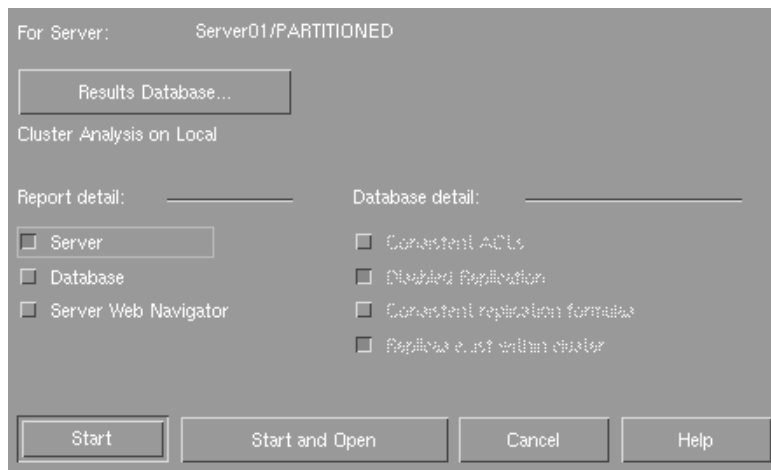
Figure 46. Cluster Analysis - Defining Results Database and Details

7. Click **Start** to run the analysis; or click **Start and Open** to run the analysis and then immediately open the Results Database to view the results.

   Figure 47 on page 117 shows the results database after it has been generated.

Cluster Analysis Results

	Source Server	Test	Result
CLUSTER			
	Server01/PARTITIONED	Number of Cluster Members	Pass
	Server01/PARTITIONED	Consistent Domain Membership	Pass
	Server01/PARTITIONED	Consistent Protocols	Pass
	Server01/PARTITIONED	Required Server Tasks	Pass

Basics

Date:	20/04/98
Time:	12:11:07
Server Name:	Server01/PARTITIONED
Cluster Name:	CLUSTER
Test Performed:	Number of Cluster Members
Test Result:	Pass

Source   **Destination**

Description

Description:	This cluster contains 2 members

Figure 47. Cluster Analysis - Cluster Analysis Database

5.4.2 Statistics Commands

You can view statistics provided by the cluster manager and the cluster replicator tasks using statistics commands at the server console.

5.4.2.1 Cluster Manager Statistics

The cluster manager task maintains cluster statistics about its view of the cluster and how often failover and workload balancing occur.

Type the command `show statistic server.cluster.*` on the server console to display the accumulated statistics for the cluster manager.

The following screen shows the output of this command.

Maintaining a Domino Cluster   **117**

```
> show statistic server.cluster.*
  Server.Cluster.AvailabilityIndex = 100
  Server.Cluster.AvailabilityThreshold = 0
  Server.Cluster.Member.Server01/PARTITIONED.Index = 100
  Server.Cluster.Member.ServNT2/PARTITIONED.Index = 100
  Server.Cluster.Name = CLUSTER
  Server.Cluster.OpenRedirects.FailoverByPath.Successful = 0
  Server.Cluster.OpenRedirects.FailoverByPath.Unsuccessful = 0
  Server.Cluster.OpenRedirects.Failover.Successful = 0
  Server.Cluster.OpenRedirects.Failover.Unsuccessful = 0
  Server.Cluster.OpenRedirects.LoadBalanceByPath.Successful = 0
  Server.Cluster.OpenRedirects.LoadBalanceByPath.Unsuccessful = 0
  Server.Cluster.OpenRedirects.LoadBalance.Successful = 0
  Server.Cluster.OpenRedirects.LoadBalance.Unsuccessful = 0
  Server.Cluster.OpenRequest.ClusterBusy = 0
  Server.Cluster.OpenRequest.DatabaseOutOfService = 0
  Server.Cluster.OpenRequest.LoadBalanced = 0
  Server.Cluster.PortName = *
  Server.Cluster.ProbeCount = 5776
  Server.Cluster.ProbeError = 0
  Server.Cluster.ProbeTimeout(mins) = 1
  Server.Cluster.Trans.IntervalAvgTime = 2
  Server.Cluster.Trans.IntervalInSeconds = 15
  Server.Cluster.Trans.IntervalsUsedInAvg = 5
  Server.Cluster.Trans.LastIntervalAvgTime = 0
  Server.Cluster.Trans.NormalizeValue = 3000
```

Notice that the Server.Cluster.OpenRedirects statistics show the frequency of successful failover and workload balancing results. These statistics provide information that is useful when reconfiguring the cluster or planning the configuration of additional cluster members.

5.4.2.2 Cluster Replicator Statistics

Type the command show statistic replica in the server console to display the accumulated statistics for both the cluster replicator task (CLREPL) and the standard replicator task (REPLICA).

The following screen shows the output of this command.

```
COMMAND SENT: show statistic replica
  Replica.Cluster.Docs.Added = 595
  Replica.Cluster.Docs.Deleted = 148
  Replica.Cluster.Docs.Updated = 10
  Replica.Cluster.Files.Local = 65
  Replica.Cluster.Files.Remote = 65
  Replica.Cluster.Retry.Skipped = 261
  Replica.Cluster.Retry.Waiting = 0
  Replica.Cluster.SecondsOnQueue = 9
  Replica.Cluster.SecondsOnQueue.Avg = 5
  Replica.Cluster.SecondsOnQueue.Max = 15
  Replica.Cluster.Servers = 1
  Replica.Cluster.SessionBytes.In = 76804
  Replica.Cluster.SessionBytes.Out = 13564356
  Replica.Cluster.Successful = 48
  Replica.Cluster.WorkQueueDepth = 0
  Replica.Cluster.WorkQueueDepth.Avg = 0
  Replica.Cluster.WorkQueueDepth.Max = 2
  Replica.Docs.Added = 595
  Replica.Docs.Deleted = 154
  Replica.Docs.Updated = 52
  Replica.Successful = 1218
```

Notice that the `Replica.Cluster.WorkQueueDepth` statistic shows the current number of modified databases awaiting replication. If this value is consistently greater than zero, enabling additional cluster replicator tasks may help you to decrease replication backlogs.

5.4.3 Statistics Reports

To enable statistics reporting in your system, you can run either the collector task or the statistics task.

By default, the statistics reports are stored in the statrep.nsf database.

5.4.3.1 Cluster Manager and Cluster Replicator Statistics

The cluster manager and cluster replicator statistics are placed in the Statistics report database statrep.nsf, under Statistics Reports, in the Clusters view.

Figure 48 on page 120 shows the cluster manager and cluster replicator statistics reports in the statistics database of Server01.

Cluster Report

for server: Server01/PARTITIONED *in cluster:* CLUSTER –

Session Information:

Boot ID:	7514670
Server started at:	17/04/98 16:55:49
Cluster statistics collected at:	20/04/98 11:46:20
Server Location:	
Server Administrator:	Admin/PARTITIONED

▶ Server Cluster Statistics:

▼

Replica Cluster Statistics:
Number of servers in replica cluster: 1
Number of successful cluster replications: 72
Number of unsuccessful cluster replications: 100
Number of docs added: 641
Number of docs updated: 23
Number of docs deleted: 148
Number of cluster replicas on this server: 66
Number of cluster replicas on other servers in this cluster: 66

Figure 48. Statistics Reports (Partial View)

5.5 Billing

This section describes setting up the billing on a server, a feature available when you install the Domino Enterprise Server license.

There are two components to billing:

• The Domino server, which generates the information for billing.

• The billing task, which produces the billing records from the information generated by the Domino server.

5.5.1 Domino Server and Billing

You can enable the Domino server to track specific activities for billing. Each type of activity is designated by a billing class.

You select which activities you want to track by adding classes to the BillingClass setting in the notes.ini file. A server with billing enabled is referred to as a "billing server".

5.5.1.1 Billing Classes

Domino uses seven billing classes to specify which activities to track. By default, Domino adds all seven billing classes to the BillingClass setting in the notes.ini file. You can modify the classes in the BillingClass setting to select the activities you want to track.

The billing classes available are:

- **Session** - Tracks when a server session starts or ends, and when any ongoing activities, document editing, or replication occur during the session.
- **Database** - Tracks when users or servers open and close a database, and the duration of use.
- **Document** - Tracks read and write activity for specified documents only.
- **Replication** - Tracks when a billing server initiates a replication with another server or a workstation.
- **Mail** - Tracks when a billing server mail router transfers messages to another Domino server.
- **Agent** - Tracks when users or servers run a scheduled agent on a server, and the elapsed runtime of agents on a billing server.
- **HttpRequest** - Tracks Web server requests.

5.5.2 Billing Task

The billing task collects the billing information provided by the billing-enabled server tasks. The Domino server places this information in the billing message queue. The billing server task then periodically polls the message queue and removes the billing records to a database or file you specify.

You can use the billing task that comes with Domino, or create your own server task using the API. For more information, refer to *Lotus Notes API 4.5 User Guide* and *Lotus Notes API 4.5 Reference*.

5.5.2.1 Enabling the Billing Task

When you install the Domino Enterprise Server, the billing task is added to the ServerTasks in the notes.ini file.

If you need to enable or disable billing, you can do the following:

- Edit the notes.ini file to add or remove the billing task to the ServerTasks stanza; and then restart the Domino server.

  For example:

  ```
  ServerTasks=Router,AdminP,Cldbdir,Clrepl,AgentMgr,Billing
  ```

- Load or unload the billing task using the `load billing` or `tell billing quit` commands at the server console.

5.5.2.2 Enabling Multiple Billing Tasks

When your server shows peak periods of user activities, you can start two or more billing tasks to manage the extra workload. Multiple billing tasks read the information from the billing queue simultaneously.

To enable multiple billing tasks, you can:

- Edit the notes.ini file to add more billing tasks to the ServerTasks stanza.

  For example:

  ```
  ServerTasks=Router,AdminP,Cldbdir,Clrepl,AgentMgr,Billing,Billing,Billing
  ```

- Load multiple billing tasks using `load billing` at the server console.

5.5.2.3 Specifying Billing Logging Frequency

The BillingAddinWakeup setting determines how often the billing task retrieves billing records from the message queue. The billing task runs for the amount of time specified and then stops. The default is 60 seconds.

To specify the billing logging frequency, you can:

- Edit the notes.ini file to add BillingAddinWakeup to the ServerTasks stanza.

  For example, to specify a frequency of five minutes (300 seconds):

  ```
  BillingAddinWakeup=300
  ```

5.5.2.4 Specifying Billing Run Time

The BillingAddinRuntime setting specifies how long the billing task runs. The default is 10 seconds.

To specify the billing logging frequency, you can:

- Edit the notes.ini file to add BillingAddinRuntime to the ServerTasks stanza.

  For example, to specify a frequency of 300 seconds:

  ```
  BillingAddinRuntime=300
  ```

5.5.2.5 Specifying Billing Output

The BillingAddinOutput setting specifies whether you want the billing task to use a database or a binary file to store billing records. The possible values are:

1 - Redirects records to billing database billing.nsf

8 - Redirects records to binary file billing.nbf

9 - Redirects records to both billing.nsf and billing.nbf

To specify the billing logging frequency you can:

- Edit the notes.ini file to add BillingAddinOutput to the ServerTasks stanza.

  For example:

  ```
  BillingAddinOutput=1
  ```

The following screen shows a notes.ini file with the SeverTasks stanza for loading the billing task, and the BillingClass stanza enabling six billing classes, and the BillingAddinOutput to the ServerTasks stanza for database billing.

```
server01:/notes/data> cat notes.ini
...
ServerTasks=Router,Replica,Update,Stats,Amgr,Sched,CalConn,AdminP,Report,Event,Billing,
HTTP,NNTP,POP3,IMAP,LDAP,Cldbdir,Clrepl
ServerTasksAt1=Catalog,Design
ServerTasksAt2=UpdAll,Object Collect mailobj.nsf
ServerTasksAt5=Statlog
BillingClass=Agent,Database,Document,Mail,Replication,Session
BillingAddinOutput=1
...
```

5.5.2.6 Viewing the Billing Database

Open the billing database billing.nsf of any Domino database. Figure 49 shows the entries for replication billing.

Figure 49. Billing Database - Replication Billing (Partial View)

5.6 Backup and Recovery in a Cluster

This section describes backup and recovery solutions when working with large amounts of data on AIX systems and Domino clusters.

If you are using Domino clustering, we assume you also require high performance, high availability and 24-hour service. Even when you have cluster replicas of all databases, you must also make regular backups to protect against data loss. When you make backups, you must bring the Domino server down before starting the backups in order to ensure data integrity. The ADSM backup system, however, is an exception to this. ADSM can make a backup on a running system and back up changes as well.

ADSM, however, puts the load on the system because it gets data from the server rather than from the files.

Imagine an installation with 1,000 users with a 40MB mailbox each = 40GB. The challenge is to make a plan backing up data and system with a high service level for the users.

When configuring backup procedures, there are basically three types of problems you want protection against:

- System failure - a system error causes loss of part or all of the data
- Application error - some data is lost
- User error - a user deletes some document by mistake

To restore from a system failure, you need a total backup. You must recover databases from the backup, not from Domino replication, because it would take too long to recover a server from a replica on other servers, and it would overload other servers. For a total backup we recommend a Domino full data backup, and eventually an AIX mksysb backup including the Domino binaries. We recommend that you make a mksysb whenever changes are made to the system. You should save a number of mksysb copies (you should have a spare copy in case of media error). A full data backup should be taken as often as possible. The longer since the last full backup, the longer it would take to recover.

But the data backup is a challenge to make. You can choose your backup device depending on the size of a total data backup. For example, a 5GB 8mm tape unit can write 1MB/second, a 7205 tape drive can write 5MB/second and a 3590 tape drive can write approximately 13MB/second in compressed mode. When you know your data size, and how long you can keep the system out of service, you can select a backup device to meet your requirements.

You can also use ADSM to back up Domino databases. ADSM has the advantage that it can back up changes, so if you have big databases, but only few updates and new documents, ADSM would be a good solution. For details, refer to *Using ADSM to Back Up Lotus Notes*, SG24-4534.

To maintain service after a system failure with data loss, you would need other working replicas in your network. Such a replica is also important for your restore process. With it, you will be able to maintain user service while restoring, and you will be able to synchronize the system up to the last change. However, you should have a new backup to avoid a load on other servers in the network.

Another possibility for backing data up when the time window is short is to use mirrored disks and then, during backup, to provide service with one disk set while backing up from the other. This method requires very good administration skills. Also, bringing the disk mirror back to service has some performance impact until data is synchronized, so this method should only be used if no other possibilities are available.

To restore data after an application error or a user error can be more complicated. If possible, prevent the loss of data by using replication, so you can get the documents back from another server. Data recovery normally requires manual restore of the missing documents. If only few documents are missing you should not restore back to the level of your last backup, but more likely you should restore the databases as a copy, and then find the missing documents and copy them back to the database.

When you copy documents back into the database, they are not exact copies. The document ID is changed, which could cause you trouble if Domino applications save and use the document number as a reference key.

5.6.1 How to Back Up an AIX Server

The two most used AIX commands for tape backup are *backup* and *tar*. We use `backup` for our examples.

To back up the notes data onto tape unit /dev/rmt0 relative to the root directory, use the following commands:

```
#cd /

#find ./notes/data -print | backup -iqf /dev/rmt0
```

The reason for backing up using relative paths (./notes...) is that this enables us to restore data into another directory later.

To back up a Notes binaries onto tape unit /dev/rmt0:

```
#find /usr/lpp/lotus /opt -print | backup -iqf /dev/rmt0
```

To make a bootable system backup we use the command `mksysb`, again to device /dev/rmt0.

#mksysb /dev/rmt0

The backup can also be made by using the SMIT interface. SMIT provides menus for system backups and data backups. Figure 50 on page 127 shows the SMIT menu for making a total backup of the Notes data directory.

```
Backup a File System

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                                  [Entry Fields]
  This option will perform a backup by inode.
* FILESYSTEM to backup                            [/notes]
* Backup DEVICE                                   [/dev/rmt0]
  Backup LEVEL (0 for a full backup)              [0]
  RECORD backup in /etc/dumpdates?                 no




F1=Help            F2=Refresh         F3=Cancel          F4=List
Esc+5=Reset        F6=Command         F7=Edit            F8=Image
F9=Shell           F10=Exit           Enter=Do
```

Figure 50. Smit Backup Filesystem Screen

AIX tape devices allow you to put more backup in serial onto one tape. To do this, you must use other tape subdevices. Device /dev/rmt* has the following subdevices:

Table 12. AIX Tape Devices

Tape Device	Rewind on Close	Retension on Open	Density
dev/rmt*	Yes**	No	High
dev/rmt*.1	No	No	High
dev/rmt*.2	Yes	Yes	High
dev/rmt*.3	No	Yes	High
dev/rmt*.4	Yes	No	Low
dev/rmt*.5	No	No	Low
dev/rmt*.6	Yes	Yes	Low
dev/rmt*.7	No	Yes	Low

To position into a specific data set, use the tctl command:

tctl [-Bnv] [-b Blocks] [-p Num] [-f Device] Subcommand [Count]

Valid subcommands are:

    weof

    eof

    fsf

    bsf

    fsr

    bsr

    rewind

    offline

    rewoffl

    erase

    retension

    read

    write

    reset

    status

If you want to position to data set number 6 and restore data, the procedure is: Go to the start of the tape, forward skip file set 5 times, and then restore. Remember to use device rmt*.1 or rmt*.5 when skipping and reading, otherwise you will get an automatic rewind or retention and lose the position.

```
#tctl -f /dev/rmt0.1 rewind
#tctl -f /dev/rmt0.1 fsf 5
#find ./notes/data -print | backup -iqf /dev/rmt0.1
```

See also *Lotus Domino Server Release 4.5 on AIX Systems,* SG24-4694, for a description on backup and restore.

5.6.2 How to Restore Files on an AIX Server

We assume the server is running and we only need to restore Domino. The Domino restore is very simple because Domino places all files in two directory trees, one for binaries, and one for data. There are no files outside these two directories (you can split your data over more directories). The first

step in the restore procedure is to restore the binaries, and then the data files. If the backup was taken relative to a directory, you can restore data into a shadow directory. If your backup was taken with full path as `/notes/data,` you can only restore back to /notes/data. The AIX command to restore files is:

```
#restore -xvqf /dev/rmt0

or

#restore -xvqf /dev/rmt0 [./notes/data/...] if only some files is restored.
```

Before you restore files it is preferable to bring the Domino server down. If this is not possible because users need the system available, you still can restore the files and then run `fixup` to make the database accessible to the server.

If you are using shared mail, you must pay special attention if you want to restore deleted mail.

Chapter 6. Adding Domino Clustering and Partitioning

This chapter describes how to update a Domino server that has not installed the clustering and partitioning binaries for cluster and partitioned server configurations.

Before you update your Domino server, we recommend that you create a backup of the Domino binaries and Domino databases.

All the information supplied in this chapter relates to Domino 4.6.1. If you are running a previous version, we recommend that you first migrate to Domino 4.6.1.

The steps you need to do are different for partitioned Domino servers and a Domino cluster. We provide the necessary steps for both situations, and several ways to implement each.

6.1 Reinstalling with Clustering and Partitioning Option

If you need to go to a Domino cluster configuration, you can follow the processes we provide in this section for each standalone Domino server or partitioned Domino server that you want to use in your future cluster.

Obviously, if you installed the clustering and partitioning binaries and the clustering and partitioning data files, you can go directly to the configuration of your cluster.

Here you have two different ways to update your Domino server for clustering purposes.

6.1.1 Manually Adding Data Files

1. Create a backup of the Domino binaries and Domino database file systems. Do not go through the following steps without a backup.

2. Stop the Domino server. Log off and log on to the system again as user root.

3. Reinstall the Domino binaries, but this time select the Domino Advanced Services option.

   If you have already installed the clustering and partitioning features, go directly to step 4 on page 132.

   If you want to configure a Domino cluster, you need the cluster programs that come with the clustering and partitioning binaries.

For the binaries reinstallation process, follow topic 3.1.5, "Installation of the Domino Binaries" on page 32.

If you had installed additional binaries from other software vendors in your Domino binaries file system, check that they are still there. Restore them from the previous backup if needed.

4. Copy the templates cldbdir.ntf and billing.ntf in the data directory of your Domino server (for example, /notes/data) from a Domino server with the clustering and partitioning data files installed. Set the owner, group and permissions of these new templates as they are for the other templates.

   These are the only additional templates that you get when you install the clustering and partitioning data files.

   If you do not have any Domino server with clustering and partitioning data files installed, just do a new Domino server installation (in another system or in the same where you are updating your Domino server, but in this case, create new file systems for binaries and data) and install the features Domino Partitioned Server and Advanced Services Data Files. Get these two templates from the new installation. Remove this installation.

5. Add the following entries in the notes.ini file of your Domino server:

```
NPN=1
ServerTasks=Router,Replica,Update,....,AdminP,Report,Event,
Billing
BillingClass=Agent,Database,Document,Mail,Replication,Session
BillingAddinOutput=1
MailClusterFailover=1
KillProcess=1
```

Figure 51. Adding Additional Stanzas to notes.ini File for Clustering

These lines are added automatically in a normal installation when you install the clustering and partitioning data files.

6. Log on to the system as user notes and start your updated Domino server.

   Once your Domino server is started, you can add it to an existing Domino cluster or add a new one.

6.1.2 Reinstalling Data Files

1. Create a valid backup of the Domino binaries and Domino database file systems. Do not go through the following steps without a backup.

2. Stop the Domino server. Log off and log on to the system again as user root.

3. Copy your actual notes.ini file in a safe place.

4. Reinstall the Domino binaries, but this time select the Domino Advanced Services option.

   If you have already installed the clustering and partitioning features, go directly to step 3 on page 133.

   If you want to configure a Domino cluster, you need the cluster programs that come with the clustering and partitioning binaries.

   For the binaries reinstallation process, follow topic 3.1.5, "Installation of the Domino Binaries" on page 32.

   If you had installed additional binaries from other software vendors in your Domino binaries file system, check that they are still there. Restore them from the previous backup if needed.

5. After the reinstallation of the Domino binaries including clustering and partitioning, log off and log on to the system again as user notes. Issue the command `notes -u`.

6. On the Domino setup window (Figure 4 on page 36), select the Domino Partitioned Server and Advanced Services Data Files options. Click the **Install** button. The setup process informs you that your previous notes.ini file is moved to another file. Write the new name of the file. We already copied the previous notes.ini file in a safe place, so we have two copies of it.

7. When the reinstallation of the data files (including clustering and partitioning) finishes successfully, click the **Quit** button and exit immediately from all windows. Stop the Notes workstation.

8. Restore the previous notes.ini file from one of the two copies adding the lines that you can see in Figure 51 on page 132.

   If you had installed additional data files (customizations, and so forth), check that they are still there. Restore them from the previous backup if needed.

9. Log on to the system as user notes and start your updated Domino server.

Once your Domino server is started, you can add it to an existing Domino cluster or add a new one.

Even though the two procedures are perfectly valid, we recommend that you reinstall the Domino binaries with clustering and partitioning and reinstall the data files with clustering and partitioning as well.

6.2 Preparing for Domino Partitioned Servers

This section describes the steps for going from a standalone server to partitioned Domino servers.

If you already installed the Domino Enterprise Server binaries during the installation of your Domino server, you do not need to do any additional procedure to configure and run partitioned Domino servers in your system.

6.2.1 Migrating a Standalone Domino Server

We assume that you have a standalone Domino server up and running. We assume that the Domino Enterprise Server is not installed.

You have two options to configure and run Domino partitioned servers in your system:

- Configure and run partitioned servers without installing the clustering and partitioning binaries.
- Configure and run partitioned servers after you install the clustering and partitioning binaries

6.2.1.1 Configuring Without Installing Partitioning Binaries

If you choose not to install the clustering and partitioning binaries, you will be able to run partitioned servers only by adding the .sgf.notespartition file to the already installed server and to each additional server you install using a different user ID.

The steps to configure the already installed Domino server are:

1. Stop the Domino server that is running.

2. Log in to your system as the notes user of the first Domino server, for example as *notes1*.

3. On the data directory of the first Domino server, /notes/data, manually create the file .sgf.notespartition. Add only one line containing only a number 1.

4. Configure the Domino server as a partitioned server, as explained in 3.2.2, "Installation Steps for AIX" on page 50.

5. Log off from the system.

6. Restart the Domino server, which now is the first partitioned server.

The steps to install and configure the second Domino partitioned server are:

1. Install the second Domino server using the already available binaries.

2. Log in to your system as the notes user of the second Domino server, for example as *notes2*.

3. On the data directory of the second Domino server, /notes2/data, manually create the file .sgf.notespartition. Add only one line containing only a number 2.

4. Configure the second Domino server as partitioned server, as explained in 3.2.2, "Installation Steps for AIX" on page 50.

5. Log off from the system.

6. Start the second Domino partitioned server.

Repeat the same process for each new partitioned server.

6.2.1.2 Configuring the Partitioning Binaries

If you choose to install the clustering and partitioning binaries, you do not need to manually configure the .sgf.notespartition file. In this case the steps you need to do are:

1. Stop the Domino server that is running.

2. Reinstall the Domino binaries selecting the clustering and partitioning binaries. All the binaries will be replaced.

3. Install the new Domino server as partitioned servers, as explained in 3.2.2, "Installation Steps for AIX" on page 50.

4. Configure the new Domino server as partitioned servers, as explained in 3.2.2, "Installation Steps for AIX" on page 50.

5. Log off from the system.

6. Restart the Domino server, which now is the first partitioned server.

Repeat the process for each new partitioned server.

6.2.2 Migrating Multiple Standalone Domino Servers

The steps for going from multiple standalone Domino servers to partitioned servers is quite similar to those for going from single standalone Domino server.

The main difference is that you have to move all the data files and users to the system that is going to run the partitioned servers.

The steps are:

1. Install clustering and partitioning in the system that is going to run the partitioned servers.

2. Move each Domino data file to the new system. Check that each server has its own user and file system.

3. Add and configure manually for each partitioned server the file .sgf.notespartition located in the Domino data directory. Make sure that each partitioned server has a different partition number.

4. Update the file .sgf.lastpartition located in the directory /usr/lpp/lotus/notes/latest/common/data/, with the greatest number you used for partition number.

5. Login as Domino user in different AIX sessions, and start each of the partitioned servers.

Appendix A. Resources

This appendix contains simple scripts and programs that you can use in your Domino environment to monitor and control your Domino servers.

A list of useful Web sites for finding information related to all the Lotus Notes Domino features is provided as well.

The information herein is provided as is.

A.1 Scripts and Programs

The scripts and programs that we give you in this section have been simplified to make them easier to understand. You can use them as a base from which to build more complex scripts and programs according to your needs.

```
#!/bin/ksh
#
#Shell script to start of Lotus Notes server on AIX
#040997  Ole Conradsen IBM danmark A/S
#

#set minimal PATH for optimized performance
PATH=/opt/lotus/bin:/notes/data:/usr/bin:/etc:/usr/sbin:/sbin:
export PATH

#set LANG if the local language is not upported
#LANG=en_US
#export LANG

#export SHARED_DPOOLSIZE= (Physical memory /2)7 = 76695844 for 1 GB ram
export SHARED_DPOOLSIZE=76695844
#enable core dump in case of a system halt then we get a dump
export DEBUG_ENABLE_CORE=1
#seach order for name lookup
nsorder=local,bind,nis

#Check if the user is notes, to prevent to start as user root.
if  [ $LOGNAME != "notes" ]
then
        echo "You must start the Domino server as user notes ! "
        echo ""
        exit 1
fi

#save 2 versions of old logs
mv /notes/data/notesserver.log.1 /notes/data/notesserver.log.2 2>/dev/null
mv /notes/data/notesserver.log /notes/data/notesserver.log.1 2>/dev/null
nohup /opt/lotus/bin/server /notes/data > /notes/data/notesserver.log &
```

Figure 52. Shell Script - Start the Domino Server

```ksh
#!/bin/ksh
#
#Shell script to close down Lotus Domino server on AIX
#040398  Ole Conradsen IBM danmark A/S

#set PATH to ensure that we have the appropriate directories
PATH=/notes/data:/opt/lotus/bin:/usr/bin:/etc:/usr/sbin:/sbin:.:
export PATH

#set Lang if needed some languages are not supported by Domino
#LANG=en_US
#export LANG

/opt/lotus/bin/server  -q &
#wait 3 min. to let Domino terminate
sleep 180

#look for  any notes process still running if any then kill it

ps -ef | grep -v grep  | grep opt | grep notes > /dev/null
res=$?

# if there are still processes hanging, then kill them
if [ $res -eq 0 ]
then

        for i in `ps -ef | grep -v grep | grep opt \
                |grep notes | awk '{ print $2 }'`
        do
#               echo "Killing notes process "$i
                kill -9 $i
        done

#Find shared memory segments and semaphores,
#append a SPACE after "m" to handle long numbers, and remove
segments

        for i in `ipcs | grep notes |grep ^m |sed 's/^m/m /' \
                |awk '/^m / {print $2}'`
        do
#               echo "removing Shared Memory " $i
                ipcrm -m $i
        done

#the same for semaphores
        for i in `ipcs |grep notes |grep ^s |sed 's/^s/s /' \
                |awk '/^s / {print $2}'`
        do
#               echo "removing Semaphore " $i
                ipcrm -s $i
        done

fi
exit 0
```

Figure 53. Shell Script - Stop the Domino Server

To use the following program (Figure 54 on page 140), you may need to adjust your notes.ini file. In this example, we remove a lot of error controls to make the program more understandable. This program is referenced in 3.6.1, "Simple Monitoring of the Response of Your Domino Server" on page 87.

```
# Program name "check_it"
# Usage: check_it DominoServer DatabaseToCheck
#include <stdlib.h>
#include <stdio.h>
#include <time.h>
/* Notes API include files */
#include "global.h"
#include "nsfdb.h"
#include "nsfdata.h"
#include "osmisc.h"

/* Program declaration */
void main (int argc, char *argv[])
{
char    fullpath_name[20];      /* pathname of database */
 DBHANDLE  db_handle;            /* database handle */
 char    buffer[NSF_INFO_SIZE];  /* database info buffer */
 char    title[NSF_INFO_SIZE];   /* database title */
 STATUS error;                   /* error code from API calls */

 error = NotesInitExtended (argc, argv);              /* V4 */
 if (error)
   {
fprintf (stderr, "\nError initializing Notes.\n");
   exit (-1);
   }
 if (error = OSPathNetConstruct(NULL,argv[1],argv[2],fullpath_name))
   exit (-1);
 /* Open the database. */
 if (error = NSFDbOpen (fullpath_name, &db_handle))
   {
   NotesTerm();
   exit (-1);
   }
 /* Get the database title. */
 if (error = NSFDbInfoGet (db_handle, buffer))
   {
   NSFDbClose (db_handle);
   NotesTerm();
   exit (-1);
   }
NSFDbInfoParse (buffer, INFOPARSE_TITLE, title, NSF_INFO_SIZE - 1);
 /* Print the title. */
 printf ("\nThe title of the database is:\n%s\n",title);
 /* Close the database. */
 if (error = NSFDbClose (db_handle))
{
   NotesTerm();
   exit (-1);
   }
 /* Terminate Notes. */
 NotesTerm();
 /* End of main program. */
 exit (0);
}
```

Figure 54. C Program - Check a Domino Server

The following shell script (Figure 55 on page 141) monitors a set of Domino servers using the C program check_it. This script is referenced in 3.6.1, "Simple Monitoring of the Response of Your Domino Server" on page 87.

```
#!/bin/ksh
##########################################################
# - This Shell_Script is a monitor of Domino servers.
# - Uses a C program called check_it which uses Notes APIs.
# - Tries to read the specific database of each server included in the list.
# 040998        Didac Marin (ISASA)
##########################################################

# Shell_Script variables
# Util. directories.
BINDIR=/notes/bin
CFGDIR=/notes/cfg
LOGDIR=/notes/log
# Seconds between each check.
INTVL=60
# LOG file.
LOGFILE=$LOGDIR/check.log
# File which contains the list of Domino servers to be checked.
ServersList=$CFGDIR/servers.lst
# Notes database to access. The best, a database unique in each server.
DataBase="check_"
Ext=".nsf"

echo "Starting Domino Servers monitoring on:" > $LOGFILE
date >> $LOGFILE

while [ true ]
do
 for server in `cat $ServersList`
 do
  echo "Monitoring server $server on:" >> $LOGFILE
  date >> $LOGFILE
  $BINDIR/check_it $server $DataBase$server$Ext >> $LOGFILE 2>&1
  if [[ $? -ne 0 ]]
  then
   echo "Server not responding..." >> $LOGFILE
   echo "Executing recovery action..." >> $LOGFILE
#   -- Put here recovery procedures according to your needs --
  else
   echo "Server responding, sleeping until next check...." >> $LOGFILE
  fi
 done
 sleep $INTVL
done
```

Figure 55. Shell Script - Monitor a Set of Domino Servers

A.2 Internet Resources

Here is a list of Internet resources where you can find almost all the information you need related to Lotus Notes Domino servers and features.

- For general information about Lotus Notes products and features, go to:

  http://www.lotus.com

- For operating system patches of known Domino server problems with these solutions and general support, go to:

  http://orionweb.lotus.com

- For:

  - Installation/Configuration procedures
  - Lotus Notes development
  - Forums
  - FAQs
  - Software downloads
  - Contacting Iris

  about all Lotus Notes products and features, go to:

  http://www.notes.net

- For information about software tools to monitor your Domino installation at the application level, go to:

  (TIVOLI, TME 10 Module for Domino/Notes)

  http://www.tivoli.com

  (PATROL knowledge module for Lotus Notes)

  http://www.patrol.com/products/pat/km/plnds.html

- For information about NotesBench, go to:

  http://www.notesbench.org

A.3 Notes Tools and Resources for the RS/6000 SP

For information about:

- Planning the hardware
- Planning SP Domino servers
- Planning SP AIX users
- Planning disk space and configuration
- Planning networks for Domino
- Planning Domino High Availability on the RS/6000 SP
- Domino server installations on the RS/6000 SP
- Domino server configurations on the RS/6000 SP

and utilities to:

- Create the AIX users for Notes in parallel
- Create the Notes file systems in parallel
- Create/Remove paging spaces in parallel
- Start/stop Domino servers in parallel
- Configure Domino servers to run in background
- Monitor the Domino servers

go to:

http://www.rs6000.ibm.com/resource/technology/notesSPcfg

Appendix B. Notes Parameters

To track the current values assigned to these pools, type the following at the server console: `show stat database`

B.1 notes.ini Parameters

<Notes Port name>=TCP, 0, 15, 0, , 12288

> Define <Notes Port name> as a valid Domino TCP/IP port name

<Notes Port name>_TcpIpAddress=0,<IP Address>:<IP port number>

> Define the IP address and IP portnumber for a previously defined <Notes Port name>

NSF_BUFFER_POOL_SIZE value *[1024 to16,777,216 bytes]*

> NIF_POOL_SIZE - controls the maximum size of memory allocated for view indexes, users of those indexes, and temporary document storage.

NSF_BC_POOL_SIZE value *[1024 bytes 1,048,576 bytes]*

> NSF_BUFFER_POOL_SIZE - sets the maximum size of the system-wide index buffer pool.

NSF_POOL_SIZE value *[1024 bytes 1,048,576 bytes]*

> NSF_POOL_SIZE is used for data structures that are allocated for each database and each thread using a database.

NIF_POOL_SIZE value *[1024 bytes 1,048,576 bytes]*

> NSF_BC_POOL_SIZE is the buffer control pool; it is used for small "management" and miscellaneous blocks of memory.
>
> Note: All settings placed in the NOTES.INI file must be in bytes.
>
> If you set one of these parameters, you should also set the remaining three.

Server_Cluster_Default_Port=<cluster port>

> This parameter tells Domino to restrict all Domino cluster traffic to the <cluster port> and leave other ports for user communication.

Server_Availability_Threshold

>Specifies the acceptable level of system resources available to a server. By setting this value for each server in a cluster, you determine how the workload is distributed among cluster members. Valid values are 0 to 100.
>
>Domino compares this value against a server's availability index; when the availability index falls below the Server_Availability_Threshold value, the server becomes BUSY. A Server_Availability_Threshold value of zero (0) indicates a fully available state and workload balancing is disabled; a value of 100 indicates the server is BUSY (since the availability index can never be greater than 100) and the Cluster Manager then tries to redirect user requests to more available cluster members.

Server_Cluster_Default_Port

>Specifies the port used for intracluster network traffic.

Server_MaxUsers

>Sets the maximum number of users that are allowed to access a server. When this number is reached, the server state becomes MAXUSERS, and the server stops accepting new Database Open requests.
>
>Use the following values to set this variable:
>
>0 - Unlimited access to server by users
>
>any number - Restricts number of active users to the number you specify

Server_Restricted

>Enables or disables server access to a server, whether or not the server is particpating in a cluster. This setting overrides any current load balancing in effect, if any, on the server. If access is disabled, the server does not accept new Open Database requests.
>
>Use the following values to set this variable:
>
>0 - Server access is unrestricted
>
>1 - Server access is restricted for the current server session. Restarting the server clears the setting.
>
>2 - Server access is restricted persistently, even after server restarts

B.2 AIX Shell Environment Variables

Korn shell syntax: export <variable name>=value

Notes_SHARED_DPOOLSIZE=16777216

Notes_SHARED_DPOOL is an AIX environment variable. It should be set in the Notes AIX user's .profile file.

Notes_SHARED_DPOOLSIZE controls the shared memory segment size that the Notes Server will request from AIX. By default, the Notes Server makes shared memory allocation requests to the kernel at a given size based on the specific UNIX platform. The effect of not setting the variable may be the error "...insufficient memory...".

The following is a working example of how to set the variable to 16 megabytes:

export Notes_SHARED_DPOOLSIZE=16777216

DEBUG_ENABLE_CORE= 0|1

The DEBUG_ENABLE_CORE AIX environment variable enables Domino to generate an AIX core dump file in case of a Domino server crash. This parameter is only necessary if you should send a dump to IBM AIX support or Lotus Support.

DEBUGSIGCHILD=0|1

The DEBUGSIGCHILD AIX environment variable controls whether the Domino server logs information about new processes started by the server in the Domino server console.

B.3 Ways to Modify the notes.ini Settings

Before making changes to the notes.ini file, use the `Show Configuration` server command at the console to check the existing settings.

To change the settings in the notes.ini file, use one of these methods:

Using a Server Configuration Document

Use a server configuration document to specify settings for a single server, a group of servers, or all servers in a domain. Use this method whenever possible, but you cannot specify all notes.ini settings in a server configuration document.

The set/modify parameter list in the server configuration document lists a set of default settings that you can specify.

Using the Set Configuration server command at the console

The `Set Configuration` server command writes the new setting to the notes.ini file. If you previously specified the setting in any server configuration document that affects this server, the new setting is written to the server configuration document specific to the server being configured, or a new document is created if necessary.

Editing the notes.ini file directly in a text editor

We do not recommend editing the notes.ini file because it can cause file errors and impair the operation of a Domino server. Use another method to set notes.ini variables if possible.

Appendix C. Special Notices

This publication is intended to help Domino and AIX specialists to plan, install and configure highly available Lotus Domino servers. The information in this publication is not intended as the specification of any programming interfaces that are provided by Lotus Domino or HACMP. See the PUBLICATIONS section of the IBM Programming Announcement for Lotus Domino and HACMP 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.

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:

IBM ®	AIX
DB2	RS/6000
SP	

The following terms are trademarks of other companies:

C-bus is a trademark of Corollary, Inc.

Java and HotJava are trademarks of Sun Microsystems, Incorporated.

Microsoft, Windows, Windows NT, and the Windows 95 logo are trademarks or registered trademarks of Microsoft Corporation.

PC Direct is a trademark of Ziff Communications Company and is used by IBM Corporation under license.

Pentium, MMX, ProShare, LANDesk, and ActionMedia are trademarks or registered trademarks of Intel Corporation in the U.S. and other countries.

UNIX is a registered trademark in the United States and other countries licensed exclusively through X/Open Company Limited.

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

Appendix D. Related Publications

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

D.1 International Technical Support Organization Publications

For information on ordering these ITSO publications see "How to Get ITSO Redbooks" on page 155.

- *Lotus Domino Server Release 4.5 on AIX Systems: Installation, Customization and Administration*, SG24-4694
- *HACMP Enhance Scalability*, SG24-2081
- *An HACMP Cookbook*, SG24-4553

D.2 Redbooks on CD-ROMs

Redbooks are also available on CD-ROMs. **Order a subscription** and receive updates 2-4 times a year at significant savings.

CD-ROM Title	Subscription Number	Collection Kit Number
System/390 Redbooks Collection	SBOF-7201	SK2T-2177
Networking and Systems Management Redbooks Collection	SBOF-7370	SK2T-6022
Transaction Processing and Data Management Redbook	SBOF-7240	SK2T-8038
Lotus Redbooks Collection	SBOF-6899	SK2T-8039
Tivoli Redbooks Collection	SBOF-6898	SK2T-8044
AS/400 Redbooks Collection	SBOF-7270	SK2T-2849
RS/6000 Redbooks Collection (HTML, BkMgr)	SBOF-7230	SK2T-8040
RS/6000 Redbooks Collection (PostScript)	SBOF-7205	SK2T-8041
RS/6000 Redbooks Collection (PDF Format)	SBOF-8700	SK2T-8043
Application Development Redbooks Collection	SBOF-7290	SK2T-8037

D.3 Other Publications

These publications are also relevant as further information sources:

- *High Availability Cluster Multi-Processing for AIX Installation Guide*, SC23-1940

- *High Availability Cluster Multi-Processing for AIX Administration Guide*, SC23-1941
- *High Availability Cluster Multi-Processing for AIX Concepts and Facilities*, SC23-1938
- *High Availability Cluster Multi-Processing for AIX Planning Guide*, SC23-1939

How to Get ITSO Redbooks

This section explains how both customers and IBM employees can find out about ITSO redbooks, CD-ROMs, workshops, and residencies. A form for ordering books and CD-ROMs is also provided.

This information was current at the time of publication, but is continually subject to change. The latest information may be found at `http://www.redbooks.ibm.com/`.

How IBM Employees Can Get ITSO Redbooks

Employees may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **Redbooks Web Site on the World Wide Web**

  `http://w3.itso.ibm.com/`

- **PUBORDER** – to order hardcopies in the United States

- **Tools Disks**

  To get LIST3820s of redbooks, type one of the following commands:

  ```
  TOOLCAT REDPRINT
  TOOLS SENDTO EHONE4 TOOLS2 REDPRINT GET SG24xxxx PACKAGE
  TOOLS SENDTO CANVM2 TOOLS REDPRINT GET SG24xxxx PACKAGE (Canadian users only)
  ```

  To get BokkManager BOOKs of redbooks, type the following command:

  ```
  TOOLCAT REDBOOKS
  ```

  To get lists of redbooks, type the following command:

  ```
  TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET ITSOCAT TXT
  ```

  To register for information on workshops, residencies, and redbooks, type the following command:

  ```
  TOOLS SENDTO WTSCPOK TOOLS ZDISK GET ITSOREGI 1998
  ```

- **REDBOOKS Category on INEWS**

- **Online** – send orders to: USIB6FPL at IBMMAIL or DKIBMBSH at IBMMAIL

Redpieces

For information so current it is still in the process of being written, look at "Redpieces" on the Redbooks Web Site (`http://www.redbooks.ibm.com/redpieces.html`). 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.

How Customers Can Get ITSO Redbooks

Customers may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **Online Orders** – send orders to:

	IBMMAIL	Internet
In United States	usib6fpl at ibmmail	usib6fpl@ibmmail.com
In Canada	caibmbkz at ibmmail	lmannix@vnet.ibm.com
Outside North America	dkibmbsh at ibmmail	bookshop@dk.ibm.com

- **Telephone Orders**

United States (toll free)	1-800-879-2755
Canada (toll free)	1-800-IBM-4YOU

Outside North America	(long distance charges apply)
(+45) 4810-1320 - Danish	(+45) 4810-1020 - German
(+45) 4810-1420 - Dutch	(+45) 4810-1620 - Italian
(+45) 4810-1540 - English	(+45) 4810-1270 - Norwegian
(+45) 4810-1670 - Finnish	(+45) 4810-1120 - Spanish
(+45) 4810-1220 - French	(+45) 4810-1170 - Swedish

- **Mail Orders** – send orders to:

IBM Publications	IBM Publications	IBM Direct Services
Publications Customer Support	144-4th Avenue, S.W.	Sortemosevej 21
P.O. Box 29570	Calgary, Alberta T2P 3N5	DK-3450 Allerød
Raleigh, NC 27626-0570	Canada	Denmark
USA		

- **Fax** – send orders to:

United States (toll free)	1-800-445-9269
Canada	1-800-267-4455
Outside North America	(+45) 48 14 2207 (long distance charge)

- **1-800-IBM-4FAX (United States)** or **(+1) 408 256 5422 (Outside USA)** – ask for:

  Index # 4421 Abstracts of new redbooks
  Index # 4422 IBM redbooks
  Index # 4420 Redbooks for last six months

- **On the World Wide Web**

Redbooks Web Site	http://www.redbooks.ibm.com
IBM Direct Publications Catalog	http://www.elink.ibmlink.ibm.com/pbl/pbl

---

Redpieces

For information so current it is still in the process of being written, look at "Redpieces" on the Redbooks Web Site (`http://www.redbooks.ibm.com/redpieces.html`). 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.

IBM Redbook 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.

List of Abbreviations

ACL	Access Control List		*GS*	Group Services
AIX	Advanced Interactive Executive		*GSAPI*	Group Services Application Programming Interface
AMG	Adapter Membership Group		*GVG*	Global Volume Group
ANS	Abstract Notation Syntax		*hb*	heart beat
API	Application Programming Interface		*HiPS*	High Performance Switch
BIS	Boot/Install Server		*hrd*	host respond daemon
BSD	Berkeley Software Distribution		*HSD*	Hashed Shared Disk
BUMP	Bring-Up Microprocessor		*IBM*	International Business Machines Corporation
CP	Crown Prince		*IP*	Internet Protocol
CPU	Central Processing Unit		*ISB*	Intermediate Switch Board
CSS	Communication Subsystem		*ISC*	Intermediate Switch Chip
CW	Control Workstation		*ITSO*	International Technical Support Organization
DB	database		*JFS*	Journaled File System
EM	Event Management		*LAN*	Local Area Network
EMAPI	Event Management Application Programming Interface		*LCD*	liquid crystal display
			LED	light emitter diode
			LRU	last recently used
EMCDB	Event Management Configuration Database		*LSC*	Link Switch Chip
EMD	Event Manager Daemon		*LVM*	Logical Volume Manager
EPROM	Erasable Programmable Read-Only Memory		*MB*	megabytes
			MIB	Management Information Base
FIFO	first-in first-out		*MPI*	Message Passing Interface
GB	gigabytes			
GL	Group Leader		*MPL*	Message Passing Library
GPFS	General Purposes File System		*MPP*	Massive Parallel Processors

159

NIM	Network Installation Manager
NSB	Node Switch Board
NSC	Node Switch Chip
OID	object ID
ODM	Object Data Manager
PE	Parallel Environment
PID	process ID
PSSP	Parallel System Support Programs
PTC	prepare to commit
PTPE	Performance Toolbox Parallel Extensions
PTX/6000	Performance Toolbox/6000
RAM	random access memory
RCP	Remote Copy Protocol
RM	Resource Monitor
RMAPI	Resource Monitor Application Programming Interface
RPQ	Request For Product Quotation
RSI	Remote Statistics Interface
RVSD	Recoverable Virtual Shared Disk
SBS	Structure Byte String
SDR	System Data Repository

Index

Numerics

ITSO Redbook Evaluation

High Availability and Scalability with Domino Clustering and Partitioning on AIX
SG24-5163-00

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!)

_____

_____

_____

_____

High Availability and Scalability with Domino Clustering and Partitioning on AIX   **SG24-5163-00**