Netfinity *White Paper*

Paul Branch — Netfinity Technology Center

paul_branch@vnet.ibm.com

Lotus Domino Clusters Overview

Abstract

The Advanced Services option[1] of Lotus Domino server 4.5 adds server clustering, billing, and partitioning features to the Domino 4.5 server. These features address availability concerns of large Lotus Notes enterprises. This paper focuses on clustering.

[1] Advanced Services is a separately priced Domino server feature. The code is available on the Lotus Domino Server CD-ROM.

Notice

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 functional 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 paper 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.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS WITHOUT WARRANTIES OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. 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. This publication could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time.

The following terms are trademarks or registered trademarks of the International Business Machines Corporation in the United States and/or other countries.

MVS	RS/6000	AIX®
OS/2®	System/390®	Parallel Sysplex
AS/400®		

The following terms are trademarks or registered trademarks of Lotus Development Corporation in the United States and/or other countries.

Lotus	Lotus Notes	Domino

Microsoft, Windows, Windows NT, and the Windows logo are registered trademarks of Microsoft Corporation.

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.

Introduction

Domino Advanced Services are a series of enhancements available for the Domino 4.5 servers, which redefine the standard for availability and reliability for mission-critical collaborative applications. Available to all customers with Domino 4.5, Domino Advanced Services include:

- **Clustering**. Allows up to six Domino servers to be grouped into a cluster providing high availability, workload balancing, and scaleability.

- **Partitioning**. A single physical server can host up to six distinct Domino servers, while providing the same level of security and reliability as if each were hosted on a standalone server machine.

- **Detailed server usage tracking and billing**. Provides valuable system utilization information.

Clustering

In simple terms, *clustering* is the linking of a group of computers together to appear as one computer to a client. Clients treat resources on clustered servers as one, highly available resource instead of separate resources. In other words, a client can request the use of a resource such as a database without being concerned "where" the database resides or "which" server is handling the request. If a server is unavailable or too busy, that server transparently passes the request to another member in the cluster capable of handling the request for the resource. From the client's (and user's) point of view, the client simply gains access to the resource requested in a timely, reliable manner.

Today's computing world regards clustering as an important, emerging technology, but in reality this technology is not all that new. Clustering has been implemented for years[2] on different types of computer platforms, from mainframes to powerful RISC workstations.[3]

If clustering is not new, why is it getting so much attention today? One reason is that recent cluster innovations focus on the clustering of Intel-based processor servers. These servers have developed a high profile primarily because of their sheer numbers and widespread use. These Intel-based servers are becoming increasingly more powerful and can contain quantities of data which would in the past have been associated only with large mainframe computers. More importantly, many companies are using these servers for mission-critical business applications, often betting their businesses on them. Further, many companies are using Intel servers as Web servers, a relatively new area of focus. If a company's Web server is unavailable, that company loses money. Given these circumstances, developing a clustering strategy containing the following characteristics is extremely important:

- High availability—if one member of a cluster fails, other member(s) of the cluster should seamlessly assume the failing member's workload. The client's work should not be interrupted by a server failure.

- Scaleability—if the demand on a cluster's resources varies widely, it should be easy to add (or remove) a member from the cluster as needed. Easy server additions relieve unanticipated

[2] Refer to *In Search of Clusters* by Gregory F. Pfister for numerous examples of clustering implementations.
[3] IBM offerings include System/390 Parallel Sysplex, RS/6000 with HACMP, and AS/400 Opti-Connect.

workload such as heavy traffic to Web servers. Easy server removal provides administrative flexibility for scheduled or unscheduled server maintenance.

- Workload balancing—if one server is inundated with a large number of requests, the other server(s) on the cluster should dynamically assume part of the workload. This balancing should increase overall system performance.

Today's Intel-based clustering solutions have implemented some of these features, but many of the solutions have limited function (for example, providing support for only two processors). However, cluster implementations are constantly evolving and improving. One such advancement is the clustering solution available in Lotus Domino Server 4.5.

Lotus Domino Server Clustering

The Advanced Services clustering option provides an *application* clustering solution for the Notes environment. Many non-Domino cluster solutions are based on specific operating systems (e.g., MVS, NT, or UNIX), meaning all servers must run the same operating system. Domino clusters are different. They are supported on any operating system which supports Advanced Services. A Domino cluster could consist of an OS/2-based Domino server, an NT-based Domino server, an RS/6000 AIX-based Domino server, and a Netware-based Domino server, or any combination of these. Depending on the operating system mix of your Notes environment, this feature could be a significant advantage. For example, in a mixed operating system environment, clustering could be used in operating system or hardware migration efforts.

In addition to operating system independence, Notes clusters provide other advantages:

- High availability—this comes in the form of failover protection for business-critical databases and Notes servers. A failover ensures that when a user attempts to access a database on a malfunctioning server, the Notes cluster servers redirect the database request to other servers in the cluster. When a clustered Notes Domino server fails, the Notes clients accessing the server continue to have access to the database through another functioning server in the cluster.

- Workload balancing—when a Notes client requests services from heavily-used Notes server, the client's request passes automatically to other cluster servers. This ensures that workload is evenly distributed across the cluster. Workload balancing helps you achieve optimum performance for your Notes servers.

- Cluster replication—this feature, unique to Notes clusters, ensures that all changes to a database or to the cluster membership itself are immediately passed to other databases or servers in the cluster. Replicas of critical Notes databases are maintained on two or more servers, and these databases are tightly synchronized to provide high availability of information. The major difference between cluster replication and traditional Notes replication is execution. While traditional Notes replication is scheduled or performed manually, cluster replication is *event driven.* For further information on cluster replication, see the following section.

- Scaleability—implementing Notes clusters provides an easy method to expand your network. This type of clustering lets you expand your processing resources effortlessly as the number of users or the workload grows. You can spread your user accounts across groups of servers,

create multiple database replicas to maximize data availability, or move users to other Notes servers or clusters as you plan future growth.

- Hardware independence—special hardware configurations are not required to use Domino clustering. There is no requirement for special adapters, cables, or shared disk drives. In fact the servers can reside in different rooms, floors, buildings, even different sites.

  The need to easily and quickly add a server to a cluster is becoming increasingly important. You might be thinking, "If you use proper capacity planning techniques, you can predict, plan and schedule an upgrade. Why is this quick upgrade capability in Notes clusters so important?" Quite simply, in today's emerging web application server environment, the rules and disciplines of capacity planning may not always apply. Many a new website or Web application has almost "capsized" because of an unanticipated rush of Web surfers. With no real way to accurately predict the workload demand of an unidentified audience, the ability to easily and quickly add more resources (such as a server) to your Web site is imperative. Notes clusters provide you with this ability so you can easily add one or more Domino servers to a cluster or replicate heavily used Notes databases onto additional servers in the cluster.

In addition to the advantages listed previously, Lotus Domino server clusters have the following characteristics:

- Up to six Advanced Servers can be configured into one cluster.
- Any database on any of the clustered servers can be clustered two to six times. Databases can be clustered to reliably achieve a service level commitment or to support large scale database demands.
- Cluster servers can be in close proximity of one another or they can be geographically dispersed.
- Cluster servers have the option of all running the same operating system or running on different supported Domino Server Advanced Services operating systems to form a cluster. In addition to operating system independence this also translates into some platform independence.

Traditional Notes Replication vs. Cluster Replication

A major difference between a traditional group of Lotus Notes servers and a cluster of Domino servers is the method used to replicate databases. Traditional Lotus Notes gives you the ability to keep multiple replicas of the same database on multiple servers. *Replication* is an important and powerful Lotus Notes process that provides synchronization of the documents in a database that may reside on different Lotus Notes servers. Replication updates a database so that any changes eventually propagate to all other replicas of the database, ensuring that a database residing on one server is synchronized with the replicas on other servers.

Traditional Notes replication occurs in one of two ways:

- It can be triggered on a timed basis or a set schedule (as shown in the figure 1).
- It can be started manually from the server console or another server.

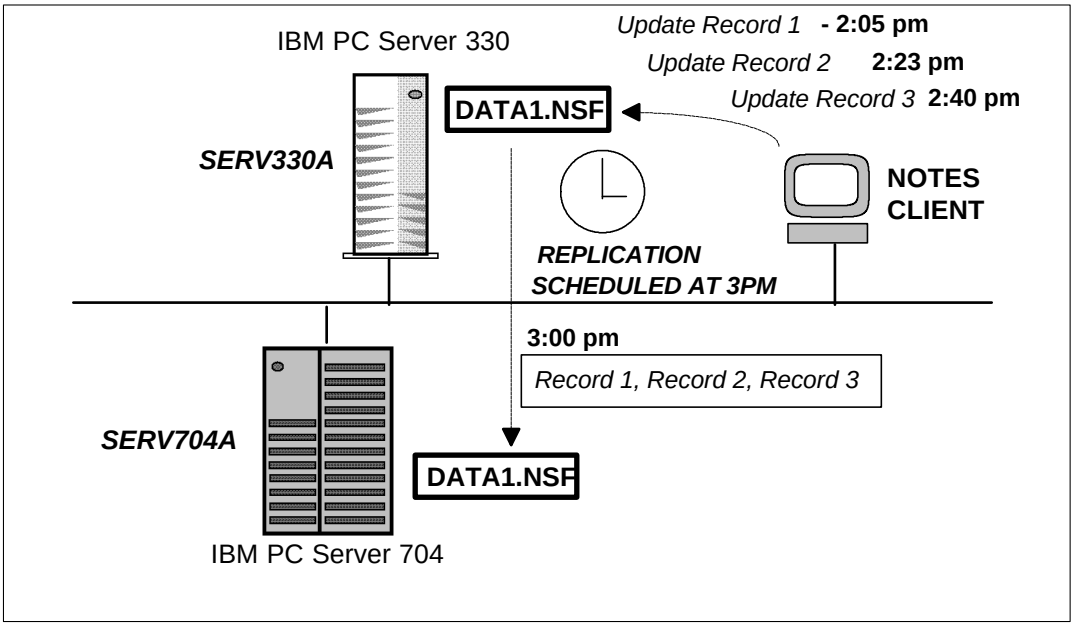

Figure 1. Traditional Notes replication

In this figure, the Notes client is updating records in the Notes database called DATA1.NSF on a Domino server named SERV330A. There is a scheduled replication every hour; in this case, the next scheduled replication is 3 p.m. When the user saves, updates or adds a record to the DATA1.NSF database, the DATA1.NSF on SERV330A is updated, but its replication copy on SERV704A is not updated *until the next scheduled replication* at 3 p.m. During this time period, the databases on the two servers are not synchronized.

How does Lotus Domino cluster replication differ from the above scenario? Since cluster replication is automatic and event driven, Notes immediately and continuously replicates the changes to other replicas of the database located in the cluster *as database changes occur*. Using this replication methodology, no matter which replica a user tries to access after a server failure, the information the user accesses is identical.

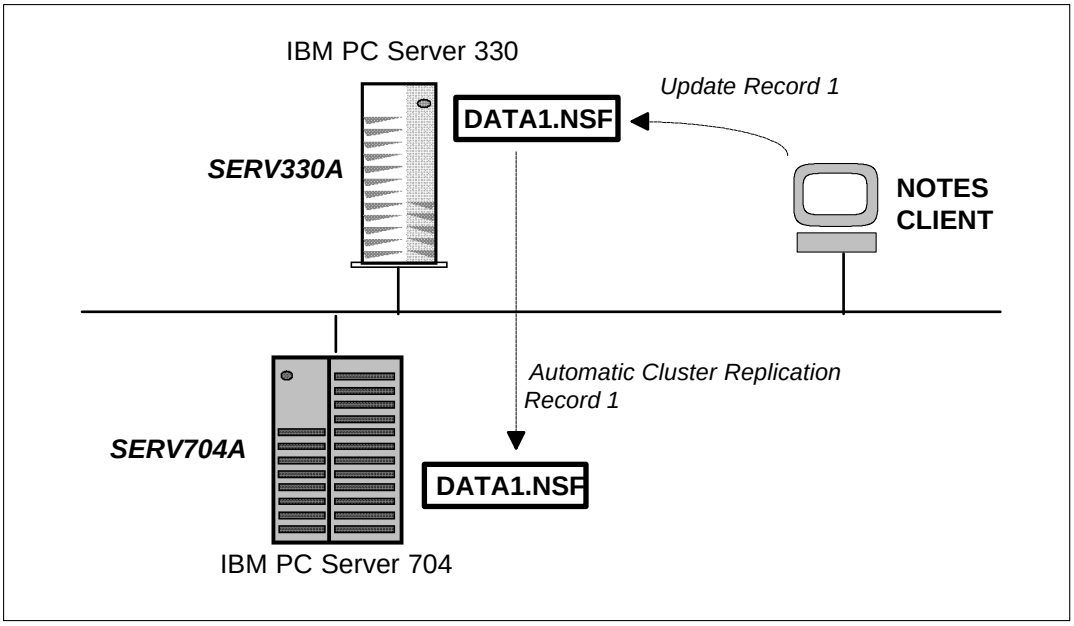

Figure 2. Domino cluster replication.

In figure 2, the Notes client is updating a record in the Notes database called DATA1.NSF on a Domino server named SERV330A. When the user saves, updates or adds a record to the DATA1.NSF database, the cluster replication takes place instantly, automatically replicating the data to the copy of DATA1.NSF on SERV704A.

How do you integrate the best elements of traditional replication and cluster replication? The best strategy is to run the cluster replication task primarily within the cluster, occasionally using traditional replication to update databases that cluster replication may have been unable to update.

Failover and Availability

A cluster's ability to redirect database requests from one server to another is called *failover*. When a user tries to access a database on a server that is unavailable or in heavy use, the user connects to a replica of the database on another server in the cluster. The switch is transparent, even though the user's client has been switched to a different server.

You do not need to install special code on Notes clients workstations to take advantage of the cluster failover feature. The client workstations only need Notes R4 (or higher) client code to handle failover, and need no special license. At this time, however, client failover features are not available to Web browser clients.[4]

When does failover happen? In general, failover occurs when a user cannot access a database in a cluster.[5] This happens either because the user cannot access the database server or because the user cannot access the database itself.

In figure 3, a user wants to open and work with the DATA1.NSF database on a Notes server called SERV330A. However, assume that SERV330A has been taken down for an operating system

[4] Web browser clients maintain their current functionality. They can still access the Domino server, but the failover function available in Notes R4 clients is not supported.

[5] For a more complete list of failover scenarios, please refer to the Lotus Notes *Network Administrator's Help* online reference.

upgrade and is unavailable. After the user issues the request, the client is redirected to an available server that has a replica of DATA1.NSF, in this case, SERV704A. The end user has the same access and database information, while being unaware that the access is to a different copy of the database on a different server.

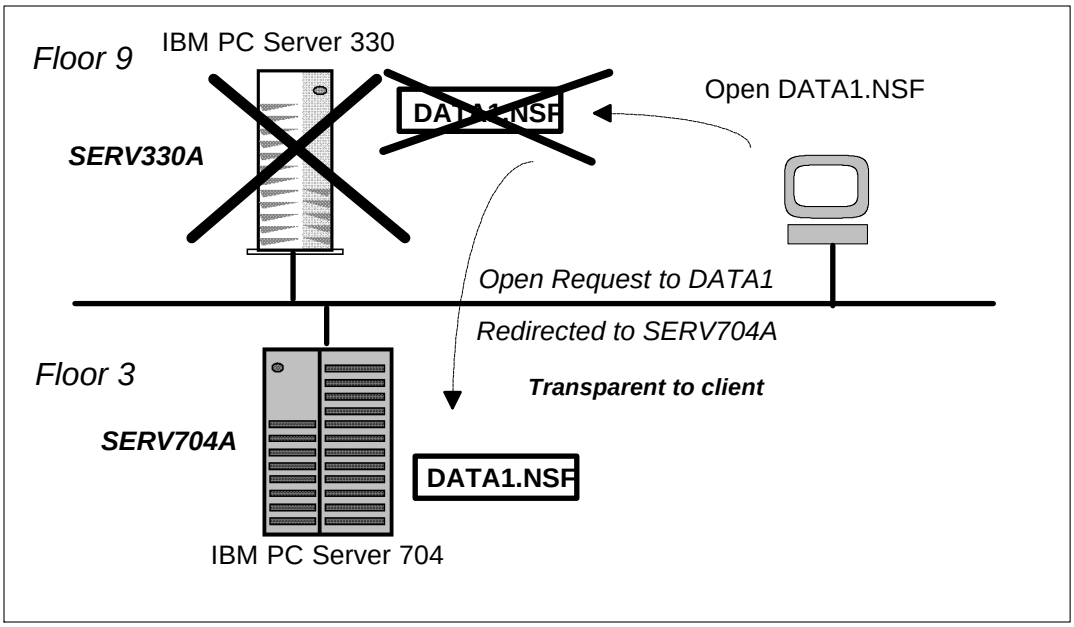

Figure 3. Domino failover.

In a non-clustered Notes environment, the client would have received a "Server not responding" message. The user would be unable to work on the database until the SERV330A server became available again.

What happens if the client is already working in a database and its server fails? This is a good news/bad news scenario. The bad news is that the failure scenario is basically the same as a non-clustered Notes environment—the client receives the "Server not responding" message and the user loses all unsaved work. However, the good news is that the user can simply exit the failing database and try to open it again. When the client reopens the failed database, failover occurs and a replica opens on a different server, allowing the user to begin working in the database immediately.

In addition to the loss of access to the server or database, failover can occur for other reasons.[6] For example, when the server is too busy, failover occurs automatically for the user—meeting the workload balancing criteria so important to effective clustering. When setting up workload balancing, you can choose from two basic algorithms, either setting a performance index threshold or setting a maximum number of users allowed on the server at one time.[7]

What does the user see when this type of failover occurs? As with the other types of failovers, the user sees nothing, except that the icons on the client's Notes workspace are stacked.

How do Notes administrators know a failover occurred? In addition to seeing a report of the event in the console log, an administrator can also view the Notes log to see if a failover or workload balance event occurred.

[6] For a complete list of failover scenarios, please refer to the Lotus Notes *Network Administrator's Help* online reference.
[7] These algorithms are defined in the NOTES.INI file, using **Server_Availability_Threshold** and **MaxUsers** parameters.

Recovering from Failures

When a failed Notes clustered server becomes available again, the administrator can simply restart the downed server. All the changes that occurred to the other servers in the cluster during the server's downtime are then replicated to the downed server. The cluster server members automatically perform these changes and re-synchronize databases. The addition of a new server to the cluster is similarly easy.

Why should this ease of use feature be so important to administrators? After all, having the computers (rather than the administrator) do the resynchronization work seems like a prerequisite for any adequate clustering solution. In reality, some other "high availability" solutions available today require complicated step-by-step manual procedures to restore configuration and production data back to its original state. Even when recovery procedures are well-documented, the potential for error increases when humans have to perform these procedures manually, and increase even more when the procedures are used infrequently. Few scenarios are worse than an unsuccessful recovery after investing time and money in what was deemed a "high availability" solution.

Planning and Administration

You can coordinate the cluster capabilities of failover and workload balancing to achieve greater database availability. However, when you try to blend these two capabilities, proper planning of the cluster is important. This section discusses different strategies you can use to optimize your Notes database and server access.

First, realize that you do not need to replicate *all* databases onto every server in the cluster. Determining how many replicas to create and where to place them requires consideration of how frequently users access a database and their need for data redundancy. For heavily used databases, you may want to create more than one replica and locate these on your most reliable computers.

In general, the more replicas of a database, the more accessible the data. But if you create too many replicas, you can add unnecessarily to the overhead of maintaining a Notes system as well as affect performance. In the figure 4, five servers belong to a Notes cluster. Note how the different databases are distributed throughout the cluster.

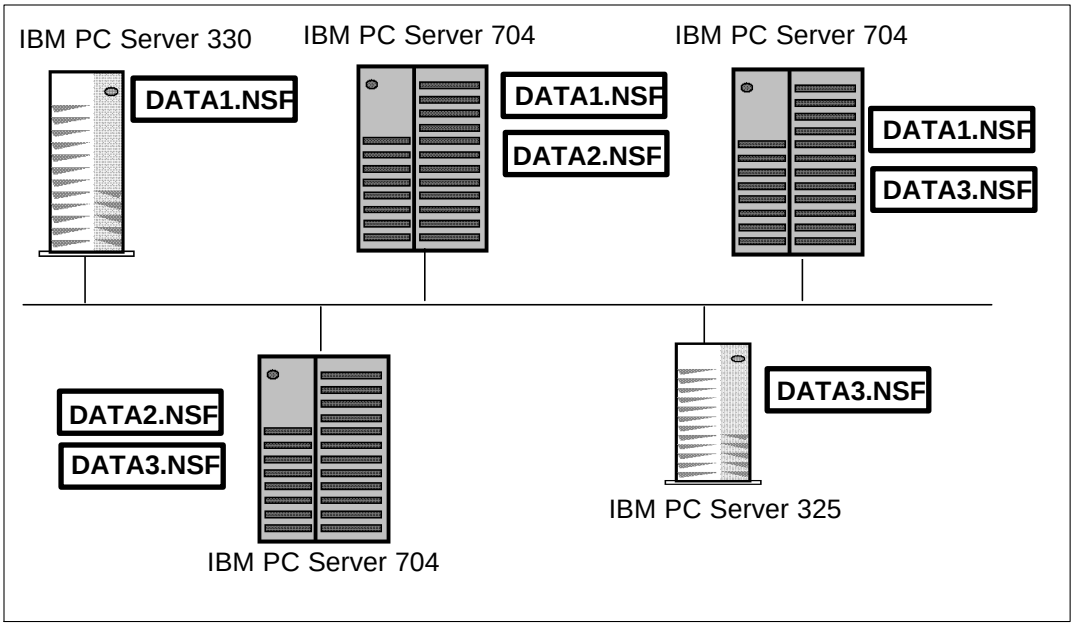

Figure 4. A typical Domino cluster.

When you distribute databases and replicas on a cluster, you should consider the following points:

- Practically speaking, having more than three replicas of a database may not provide you with significant additional availability. If users can adequately access a particular database from one or two servers only, do not increase the number of replicas in the cluster. The probability of losing both servers at the same time is minimal.

- There are situations where users require the constant availability of a specific database. In these cases, consider placing replicas on every member in the cluster—provided you have adequate disk space and resources to do this.

- Not all databases require multiple replicas in a cluster. For example, a server log does not need to have a replica on another server.

It may occur to you that placing multiple replicas of databases throughout the cluster is an inefficient use of your disk resources. You may read about some other cluster solutions that permit multiple clustered servers to access and share a group of disks (figure 5).

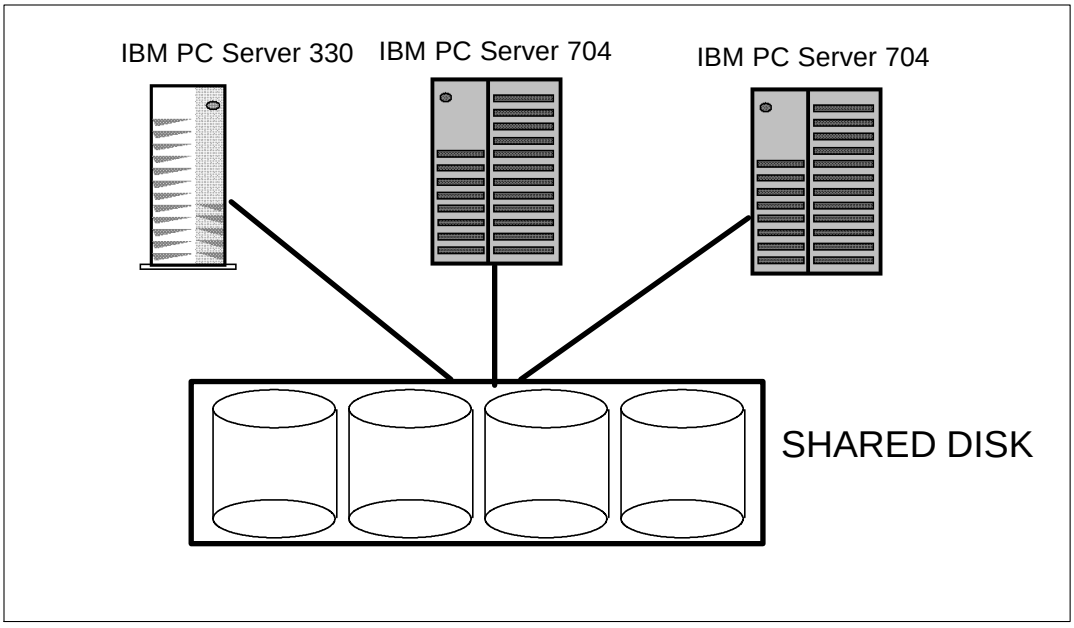

Figure 5. A shared disk cluster.

Note the single point of failure in this type of configuration. If one of the important objectives of clusters is to provide a high availability solution, and a single point of failure exists, high availability is in jeopardy. Although a true high availability solution requires additional disk space, its design greatly decreases the risk of having single point of failure. For this benefit, no specialized high availability hardware is required, unlike the hardware needs of many specialized shared disk solutions. For example, in a shared disk solution, cable lengths can only be a certain distance from the servers, greatly limiting the geography of a network. Notes clusters can be located on different floors of a building, or in different buildings, or even in different countries.

Note also the potential performance bottleneck in a shared disk construction. While these types of configurations exist today in other clustered environments, they provide adequate performance only with the aid of specialized hardware and connections. Experienced administrators realize that providing adequate disk throughput on a single server can be challenging at times. Imagine the administrative task of analyzing and balancing the disk workload with multiple clustered servers trying to access the same group of disks and data. As the number of servers participating in a cluster increases, this administrative task becomes more and more complex. Notes clusters do not require specialized disk hardware and can use traditional techniques for analyzing disk performance.

Network Traffic Considerations

Cluster replication constantly updates replicated databases over your network. Although these transactions are primarily small records, potential congestion on the network may be a performance concern for a Notes cluster administrator. If replication-generated traffic is a concern, one option is to isolate the cluster traffic by setting up a private LAN dedicated for intra-cluster communication between servers. Compared to the cost of network bottlenecks in lost productivity, the cost of this private LAN is minimal—an additional network adapter in each server, a hub, and wiring.

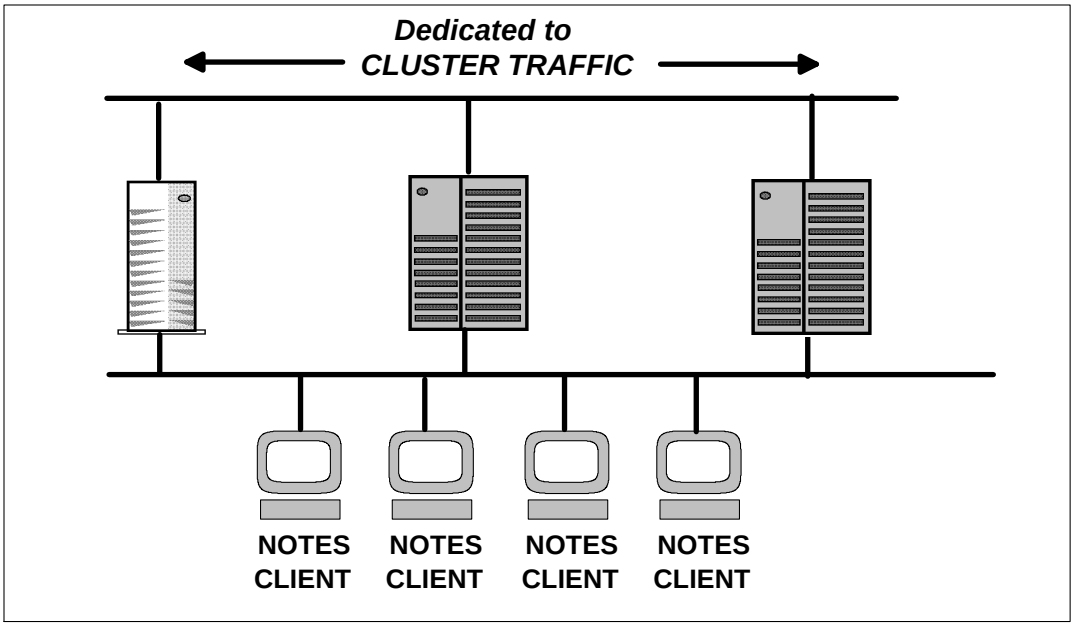

Figure 6. Intra-cluster communication.

Cluster Setup Strategies

After reviewing your Notes environment, you should plan a cluster strategy that balances your user's requirements for data availability with the physical ability of each computer in your cluster to manage additional workload. Depending upon your needs, you can choose from three major strategies for optimizing database and server access.

The first strategy is the *high failover* strategy, one that concentrates on keeping important databases highly available to the clients. One obvious type of database for which this strategy works well is a user mail database.[8] Having multiple replicas of these types of databases contained on different servers helps to ensure that users can always access the data they need.

The second strategy is *active workload balancing.* With Notes clusters' optional workload balancing feature, you can distribute the workload of heavily-used databases across multiple servers in the cluster. Mail databases are again good candidates for this type of strategy, for implementing it ensures that no single server has a significantly larger number of mail files than any other server.

The third strategy is a *combination* of the these two strategies. When you have a database that users access continuously, such as a special discussion database, you need not only high availability, but also the ability to distribute within the cluster the heavy workload created by users accessing the database.

[8] The **MailClusterFailover** setting in the NOTES.INI file enables Mail Router request failover.

Summary

In summary, using the Lotus Notes clusters feature introduced in Lotus Domino Server 4.5 provides an application clustering solution for the Lotus Notes environment. This implementation combines the important basic clustering functions of high availability, scaleability, and workload balancing into one integrated package. With high availability, you can place multiple copies of databases onto multiple servers, allowing constant user access with few service interruptions. With scaleability, you can easily add and remove servers on the cluster. And with workload balancing, you can hand the distribution of the cluster's workload over to the clustering technology, ensuring the cluster dynamically adjusts to fluctuations in user demand.

In general, Notes clusters currently allow you to group up to six servers to form a cluster. Because Notes clustering is an application clustering solution, it provides you with a degree of operating system independence. Any operating system supported by Advanced Services can participate in the cluster. In addition, you can also run different operating systems within the same cluster. Notes clustering is also easy to implement. It does not require any special hardware or rewriting of your Notes applications. If you currently have Lotus Notes 4.5, basically all you have to do is install the Advanced Services code, configure the cluster, and replicate your important databases throughout the cluster.

One major benefit of using a Lotus Domino cluster instead of traditional Lotus Notes servers is that the replication of data becomes instantaneous and continuous to all functioning servers. Traditional replication is either run manually or triggered on a timed basis, a method that can leave gaps of time where all replicas of data are not synchronized. The Notes clusters' event-driven replication method eliminates this risk.

Implementing a clustering solution involves a careful evaluation of how your network can benefit most from the clustering. Considering your users' needs as well as the physical capabilities of your equipment is critical. Cluster setup strategies can concentrate on high availability of data, on balancing workload for the servers, or on a strategy that incorporates both these concerns.

Although many new Intel-based clustering solutions are being touted in today's marketplace, many of these provide only a subset of true clustering. For example, some solutions merely provide for high availability while ignoring scaleability and workload balancing. In comparison, Notes clusters provide a comprehensive and advanced clustering solution for the Lotus Notes environment. In many ways, Lotus Domino clusters lead the Intel-based server clustering pack.