Information Management software

Availability Solutions with Informix® Dynamic Server 11

Availability Solutions with IBM Informix Dynamic Server 11.10

Madison Pruet
Ajay Gupta

The addition of Multi-node Active Cluster for High Availability (MACH11) technology significantly expands the Informix Dynamic Server (IDS) high availability options to provide increased failover, capacity, flexibility, and scalability.

Traditionally, IDS has provided multiple robust solutions for supporting high availability data replication options. Previous releases of IDS have supported two replication technologies: Enterprise Replication (ER) and High Availability Data Replication (HDR). By using these technologies together, customers are able to achieve very high levels of data availability. . Both replication technologies can be integrated with each other and coexist with other availability solutions such as disk mirroring.

While HDR and ER have been features of IDS for many years, and have proven to be highly reliable and low-maintenance technologies. IDS now adds support for two new types of secondary servers:

- Secondary servers that share the same physical disk. The Shared Disk Secondary (SDS) servers provide increased availability by allowing one or more instances of the IDS server to attach to the same disk subsystem, providing redundancy for the server in addition to data redundancy solutions
- Additional remote secondary servers. The Remote Standalone Secondary (RSS) servers extend HDR to provide multiple local or remote backup servers that also replicate the data.

Both SDS and RSS servers provide customers a way to obtain increased capacity by distributing workload across multiple servers.

Customers can choose any of these solutions on their own. They become even more powerful by combining all three types of topologies together. Adding Enterprise Replication, a completely customized availability solution can be delivered to meet each unique availability requirement.. These configurations are simple to set up and maintain, and are highly scalable.

1. 0   A Review of Existing IDS Availability Solutions

IDS has always provided robust data replication features facilitating a very high degree of availability, both during online operation and in the case of a failure. Historically, availability solutions for IDS customers have included:

- Backup and restore
- Disk mirroring
- High Availability Data Replication (HDR)
- Enterprise Replication (ER)

The characteristics of each of these solutions are as follows:

1.1 Backup and Restore

IDS was one of the first database servers to introduce online backup. With backup and restore, the customer periodically performs an online backup of the system onto an external device such as tape media. The media is then taken off-site to a secure location. In the event of a catastrophic system loss, the external backup is used to restore the system to the point at which the backup was taken.

The logical logs which contain a list of changes made to the database are also backed up. By applying the logical log backup to the system being restored, the database can be recovered up to the point at which the last backup of the logical logs was made.

1.2 Disk Mirroring

Disk mirroring is commonly performed using software or hardware to mirror the database chunks (chunks are physical storage location of data in IDS).

IDS provides built-in support for dual or mirrored chunks. When using mirrored chunks, the chunks are stored in two separate files: the primary and its mirror. Writes occur to both the primary and to the mirror chunk. If the primary chunk becomes unavailable while the server is active, then the server will automatically switch to the mirror chunk.

In recent years, customers have increasingly chosen to use hardware disk mirroring rather than mirrored chunks. With hardware disk mirroring, the entire disk complex is mirrored, usually through some form of RAID or other disk volume manager. The advent of Storage Area Network (SAN) and Network Attached Storage (NAS) solutions has made it possible to separate the physical storage media from the systems that use that storage. In effect, instead of the disk media being attached to the server, the server is attached to the disk. This means that it is possible to have a separate idle standby system with IDS installed that can be used to provide availability in the event of the loss of the main server. Since it is possible to locate the mirrored disk some distance away from the primary, a high degree of availability is possible.

1.3 High Availability Data Replication (HDR)

As early as IDS Version 7, Informix adopted HDR technology, which is fully integrated within the data server. HDR is very easy to set up and administer and does not require any additional hardware or software for automatically handling server or disk failures.

HDR maintains two identical IDS server instances on servers with similar configurations and operating systems (see Figure 1). HDR employs a log record shipping technique to transfer the logical log records from the primary server to the secondary server. The secondary server is in perpetual roll-forward mode so that data on the secondary server remains current with data on the primary server. The secondary server supports read access to data, allowing database administrators to spread workload among servers.
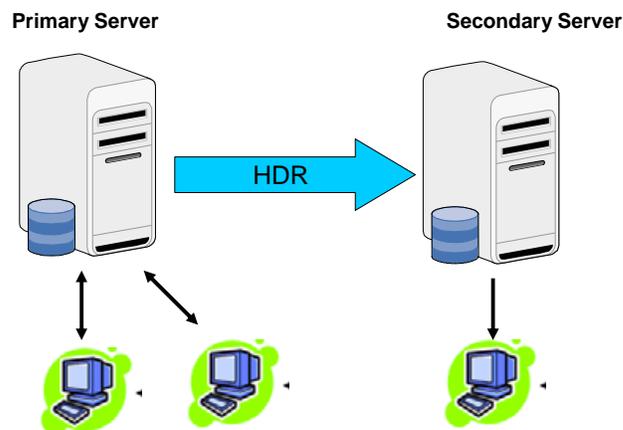
Figure 1. High Availability Data Replication (HDR)

The secondary server can be configured to operate in synchronous (SYNC) or asynchronous (ASYNC) mode. In SYNC mode, HDR guarantees that when a transaction is committed on the primary server its logs have been transmitted to the HDR secondary server. In ASYNC mode, transaction commitment on the primary and transmission of updates to the secondary are independent, providing better performance but possible risk of lost transactions.

HDR provides automatic failover to redirect client applications to the new primary server without missing a beat. With the DRAUTO parameter set, if the primary server fails, the HDR secondary server automatically takes over and switches to a standard or primary server (based on the DRAUTO value). When the original primary server becomes available, it is synchronized when HDR is restarted.

Current HDR replication technology also supports automatic client redirection. This feature makes failover transparent to the application. To activate automatic client redirection, the primary and secondary servers must be defined as a group in the SQLHOSTS file. Clients use the group name to connect to the IDS server. The network layer and the client-server protocol ensures that the client is always connected to the primary server in the group. If the primary server fails and the secondary server becomes the new primary server, clients connected to the group will be automatically connected to the new primary server. This means that end user applications will not experience any outage, even though the application is now pointing to a different database server.

1.4 Enterprise Replication (ER)

Enterprise Replication (ER) provides replication of data across multiple independent IDS servers, and has the ability to support both "active-passive" and "active-active" replication. That is, any of the servers participating in the ER cluster can accept both read and write transactions. Conflicts between servers are resolved in ER by reconciling transactions. ER can also be used to replicate individual tables or subsets of tables rather than the entire database. This is different from HDR, since HDR requires an exact replica of the data – including table and database schemas. ER is designed to support multiple servers with complex topologies.
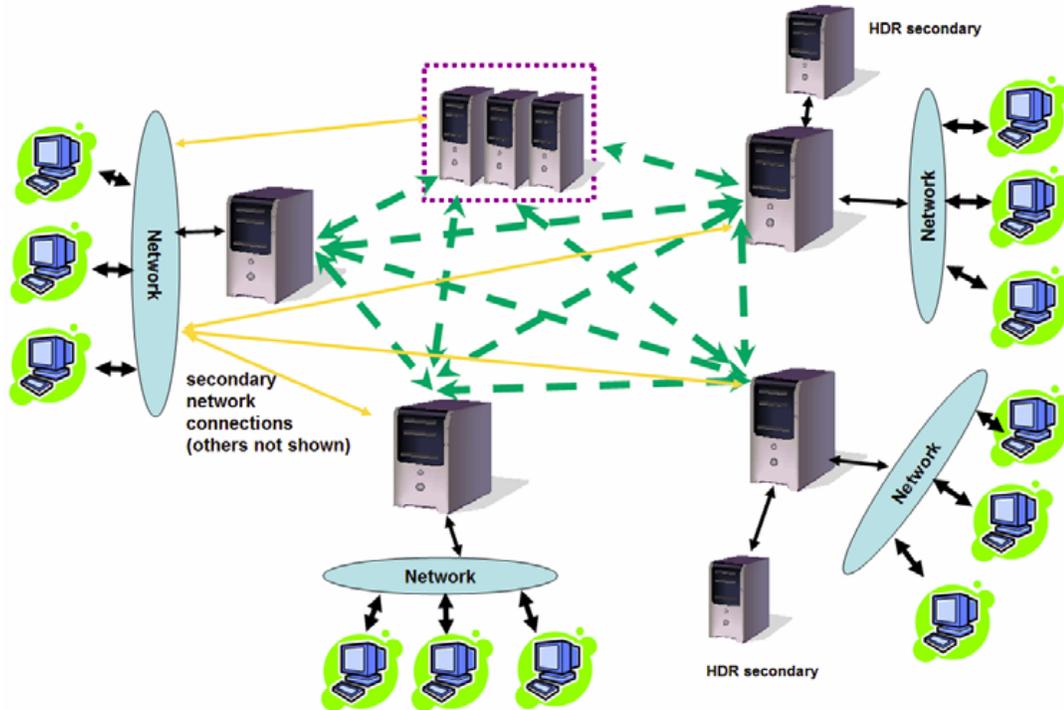
Figure 2. Enterprise Replication

ER replicates data asynchronously, which provides the ability to accommodate slow or unreliable networks and not worry about any inherent delays.

ER can be configured so only a subset of data (even using different table schemas) can reside locally. This is ideal when regional offices, for example, need to be restricted to view only their data, while an entire copy is maintained at a corporate office. The ER Redbook *(get name and link)* contains a very detailed description of the technology and design options for this functionality.

2. 0   New High Availability Solutions

IDS 11 provides two additional types of server configurations that can be used in conjunction with existing HDR and ER configurations: Remote Standalone Secondary Servers and Shared Disk Secondary Servers. Additionally, new Continuous Log Restore functionality makes it possible to manually maintain a backup system.

2. 1　Remote Standalone Secondary (RSS) Servers

The Remote Standalone Secondary servers extend HDR by allowing multiple copies of the database in both local and geographically remote locations. These secondary servers, like HDR, can be accessed by the client applications for query activates. Logical logs are continuously transmitted from the primary server and applied to the database on the RSS server.
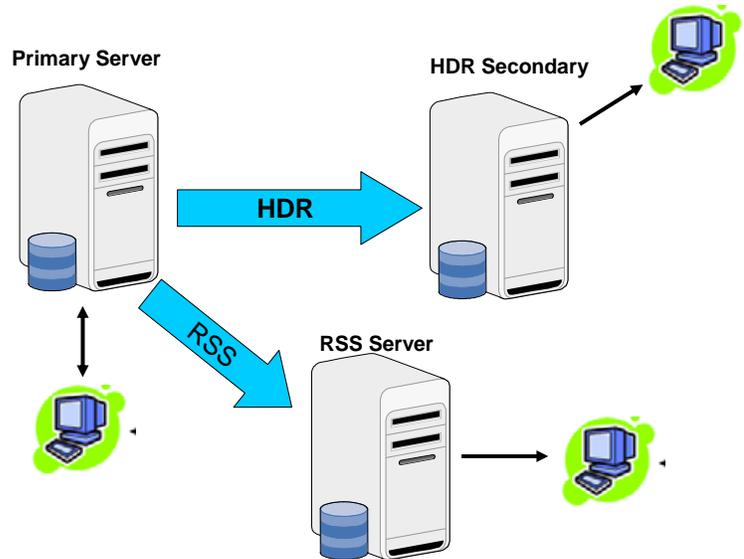
Figure 3. Remote Standalone Secondary (RSS) Server

RSS servers use a fully duplexed communication protocol, allowing the primary server to send data to the RSS servers without waiting for an acknowledgement that the data was received. Using full duplexed communication means that RSS servers have very little impact on the primary server's performance. Many RSS servers can be established, providing backup systems in remote locations around the world, and delivering data close by where it is needed.

The HDR secondary is still the server that is failover-ready from the primary. There are 3 ways that the RSS server can change roles:
1. The failover-ready HDR server becomes unavailable. In this case, one of the RSS servers can be assigned the failover-ready HDR server role.
2. The primary server becomes unavailable, and the current HDR secondary assumes the role of the primary. One of the RSS servers can then be assigned the role of the HDR secondary server.
3. If both the primary and the HDR servers become unavailable, one of the RSS servers can become the primary server. Another RSS server can then be assigned the HDR secondary role.

Multiple RSS servers in geographically diverse locations can be used to provide continuous availability and faster query access than if all users were directed to the primary server. The application traffic that is read-only can be sent to local RSS servers. For example, RSS servers can feed data to Web applications that do not require up-to-the-minute data currency. If the applications need to update the data, they can connect to primary, otherwise they read the data from the local RSS server. This configuration will reduce network traffic and the time required by the application to access the data.

In figure 3a, below, remote servers can access local database servers to minimize latency and improve performance.
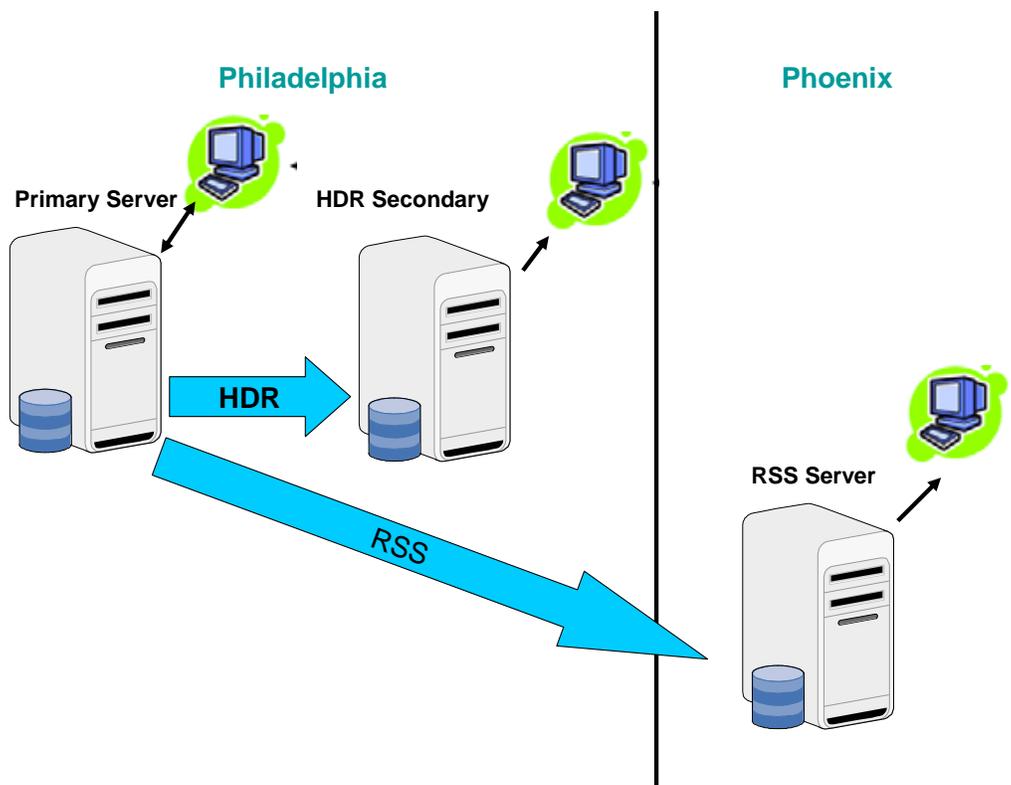
Figure 3a. Physically Remote Standalone Secondary (RSS) Server

2. 2　Shared Disk Secondary (SDS) Servers

SDS servers access the same physical disk as the primary server. They provide increased availability and scalability without the need to maintain multiple copies of the database.
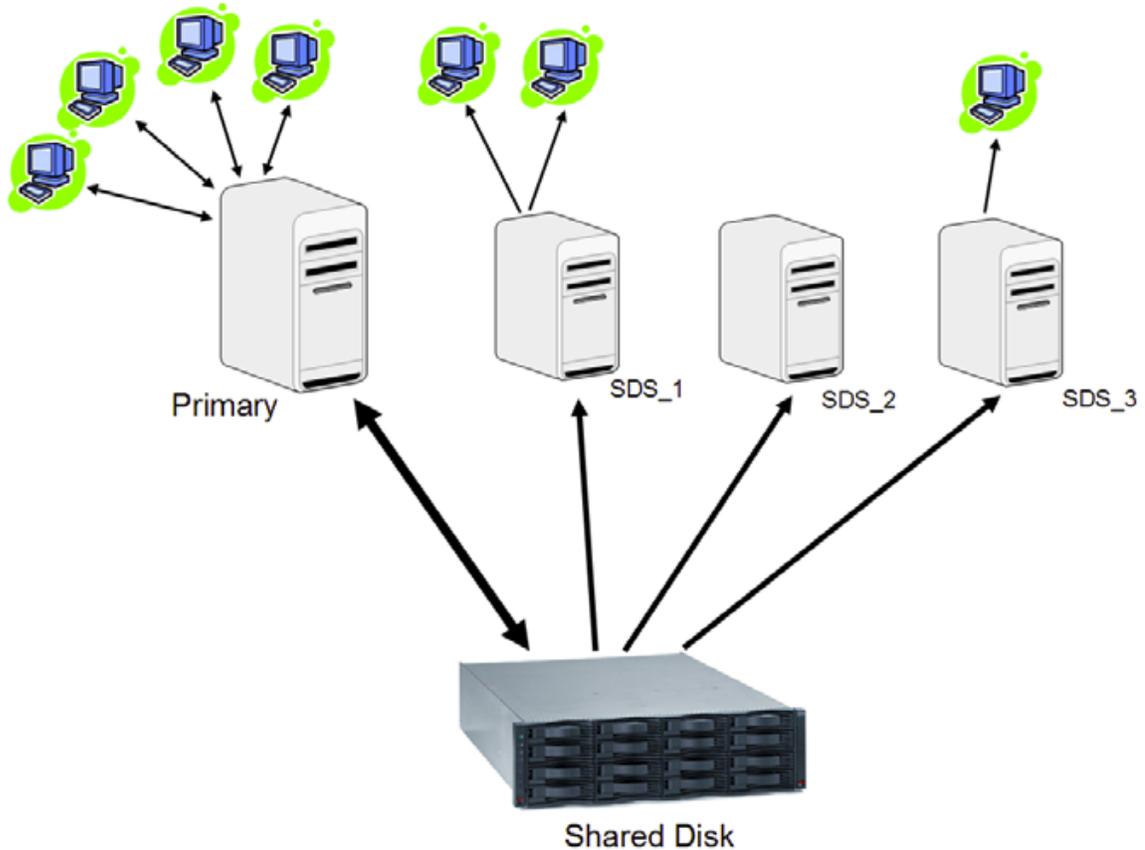
Figure 4. Shared Disk Secondary (SDS) Servers

An SDS server can be made available very quickly. Once configured, an SDS server joins an existing system and is ready for immediate use. Because SDS servers also use fully duplexed communications with the primary, having multiple SDS servers has little effect on the performance of the primary server. SDS servers are completely compatible with both hardware and software-based disk mirroring.

If the primary server becomes unavailable, failover to an SDS server is easily accomplished. The specified SDS server becomes the new primary server and all other SDS servers automatically recognize the new primary.

Multiple SDS servers also provide the opportunity to offload reporting and other functionality from the primary server. For example, a system with four SDS servers can have two allocated for analytics and two for read-only Web site data. During the holiday season, all four SDS servers could be allocated to Web site data to support the extra traffic.

2. 3　　　Continuous Log Restore

Continuous Log Restore is useful when the backup database server is required to be fairly current, but the two systems need to be completely independent of each other for reasons such as security or network availability. Continuous Log Restore can also be useful when the cost of maintaining a persistent network connection is too high. With Continuous Log Restore, log files are manually transferred to a backup database server where they are restored.

Continuous log restore is a robust way to set up a hot backup of a database server. The hot backup of the primary IDS server is maintained on the backup server, which contains similar hardware and an identical version of IDS. To configure a backup server using Continuous Log Restore, a physical backup of the primary server is created and the backup copy is transported to the backup server. The backup is then restored on the backup server. After the restore is complete, the backup server is ready for a logical recovery. In the event that a logical log on the primary server becomes full, it is backed up and then transported to the backup server where logical recovery (log roll forward) is performed. The secondary server remains in log restore suspended state after the last available log is restored. With the server in this state, another logical restore can be started immediately after additional logical logs become available. Operation of Continuous Log Restore is shown in figure 5.
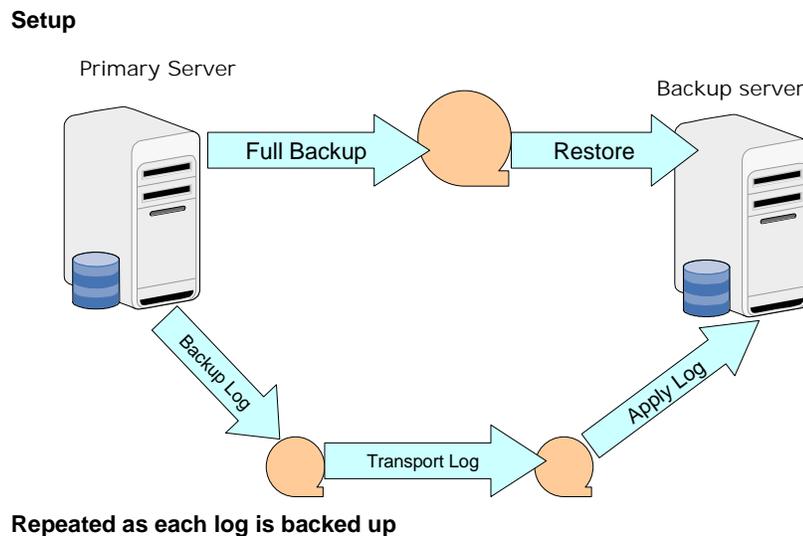
Figure 5. Continuous Log Restore

Should the primary server become unavailable, a final log recovery is performed on the backup server, which is brought up in online mode as the primary server.

Continuous Log Restore can be combined easily with the other high-availability solutions, such as shared disk and remote secondary servers, or with hardware solutions, such as cluster failover.

3. 0   The Recoverable Group and Transfer of the Primary

In previous versions of IDS the only failover option available was to failover the primary server to an HDR secondary. IDS 11 adds two new failover server configuration options. These are useful not only in a failure or disaster situation, but also for performing maintenance tasks on the database server. The configuration of servers can be thought of as a group of servers that act as a recoverable unit. This recoverable unit consists of a single primary server and one or more secondary servers.

A multi-node active cluster (MACH11) consists of a single primary server and one or more secondary servers. The secondary servers can include any combination of SDS, RSS, and HDR secondary server.

3. 1   Recoverable Groups Based on the Shared Disk Secondary
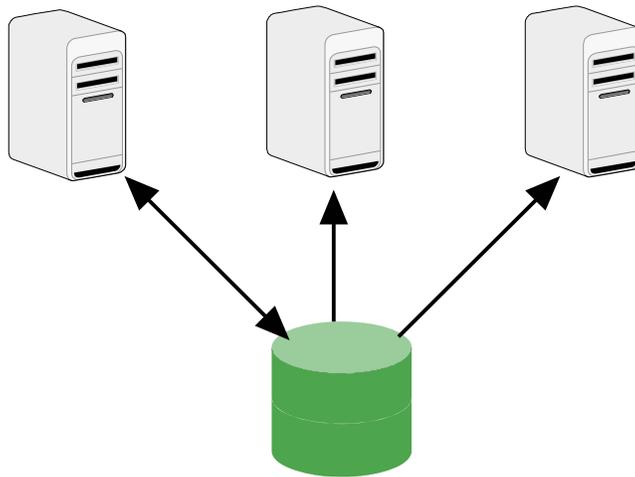
Figure 6.   Primary with two SDS servers

Figure 6 shows an example of a primary server configured with two SDS servers. In this case, the primary server role could be transferred to either of the two SDS servers. This would include the case where the primary needs to be taken out of service for a planned

11

outage, or because of a failure of the primary server. Since both of the SDS servers are reading the same disk subsystem, there is no difference in which of the two SDS servers becomes the primary; they are equally able to assume the role of the primary server if they are similarly sized servers. This is illustrated in the following figure.
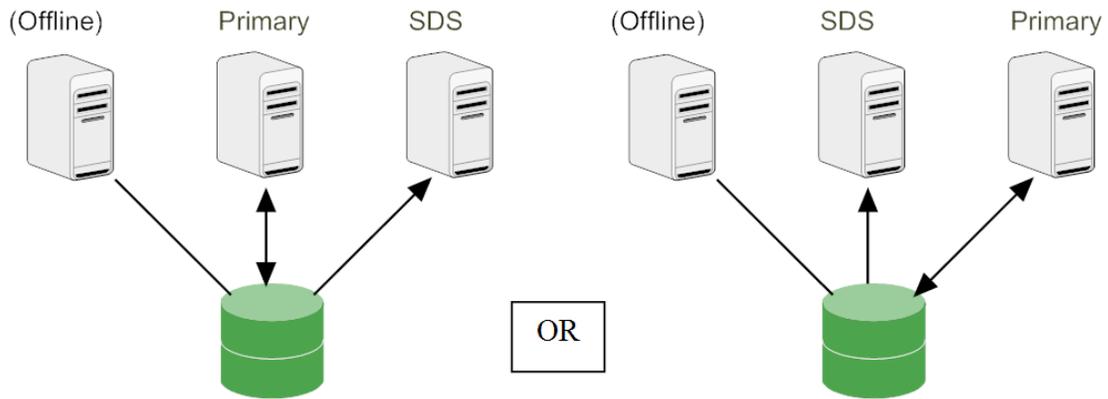
Figure 7 Transfer of the Primary

There are also a number of recovery options to handle the loss of the shared disk itself. The most common recovery option is to have the shared disk use either RAID technology (such as RAID 5), or to use disks based on SAN technology, which may include some form of remote disk mirroring. Since the disk and its mirror can be located in different areas, this provides a high degree of availability for both planned and unplanned outages of either the server(s) or of the disk subsystem(s).
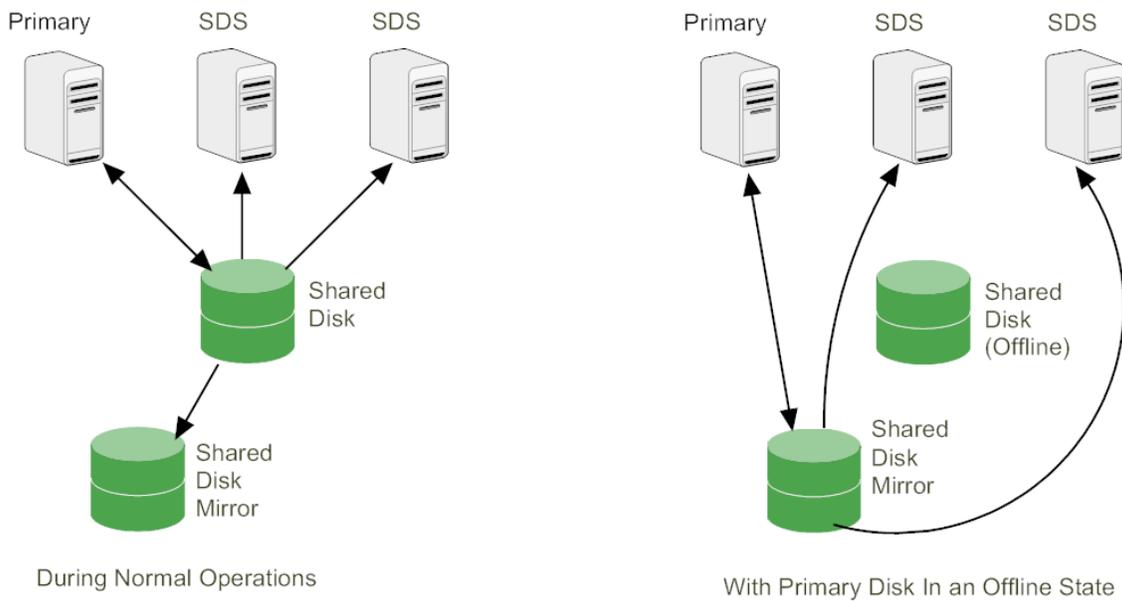
Figure 8. Disk Availability with SDS Servers

In addition to configuring a mirrored disk subsystem as in the previous example, you may want to have the extra redundancy of additional servers. For example, you may want to have the primary and two SDS servers in the figure 8 contained within a single blade server enclosure. The configuration in figure 9 is an attractive solution when you need to periodically increase processor read processing ability such as when performing large reporting tasks.
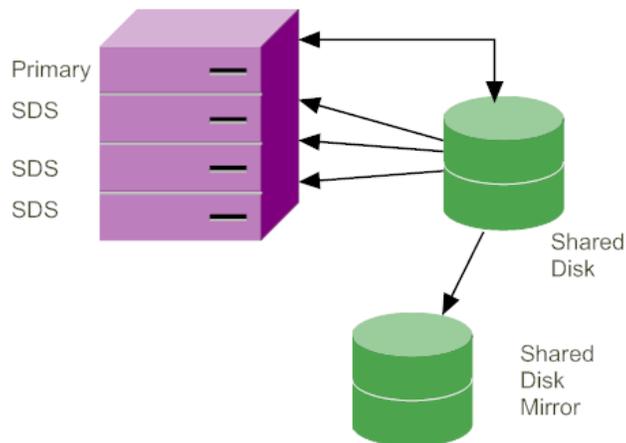
Figure 9. Single Blade Server Housing the Primary and Three SDS Servers

13

In such a configuration, you may decide to avoid the possible failure of a single blade server by using multiple blade servers, as in the following illustration:
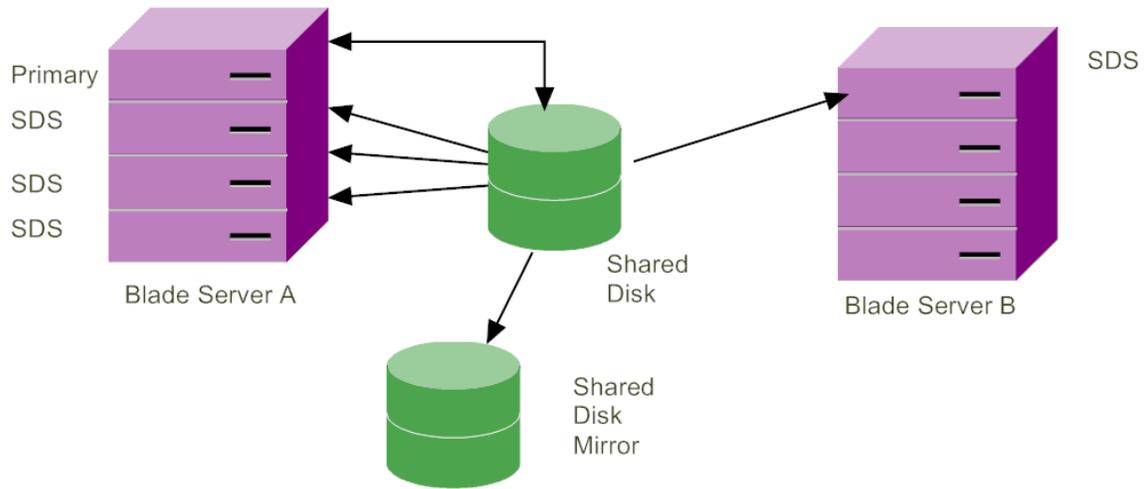
Figure 10. Two Blade Servers being used to house SDS Servers

In the previous illustration, if Blade Server A should fail, it would be possible to transfer the primary server role to the SDS server on Blade Server B. Since it is possible to bring additional SDS servers online very quickly, you can dynamically add additional SDS servers to Blade Server B, as in the following illustration.
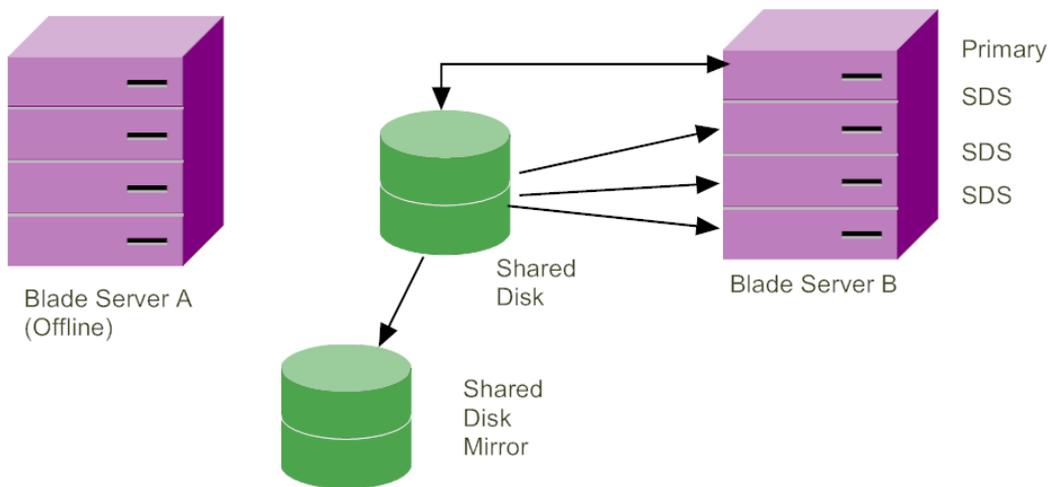
Figure 11. After transfer of primary server from Blade Server A to B
and dynamically starting additional SDS servers

Because of limits on the distance that disk mirroring can support, you may need an alternative to using shared disks and shared disk mirroring to provide availability. For example, if you prefer that there be a significant distance between the two copies of the disk subsystem, you might choose to use either an HDR secondary or an RSS server to maintain the secondary copy of the disk subsystem. The following illustration shows an example of an HDR secondary server in a blade server configuration.
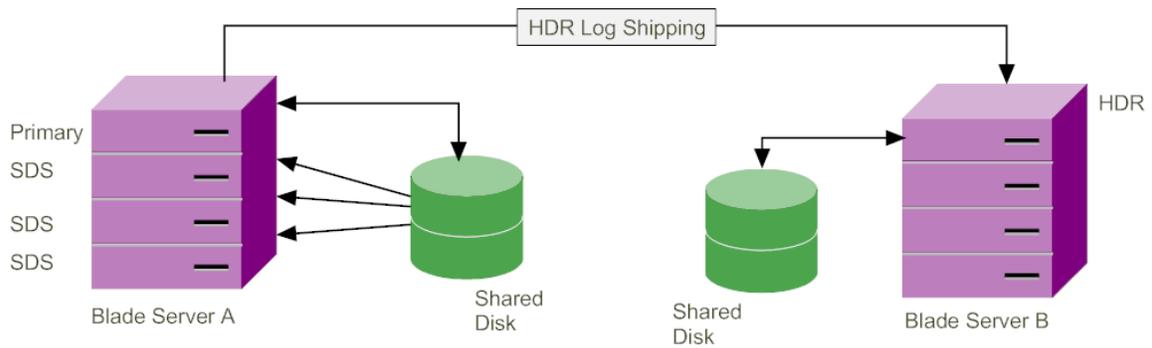
Figure 12. Using an HDR secondary in conjunction with SDS Servers

In this configuration, if the primary server should fail, but the shared disks are intact and the blade server is still functional, it is possible to transfer the primary server role from the first server in Blade Server A to another server in the same blade server. Changing the primary server would cause the source of the remote HDR secondary server to automatically reroute to the new primary server, as illustrated in the following diagram:
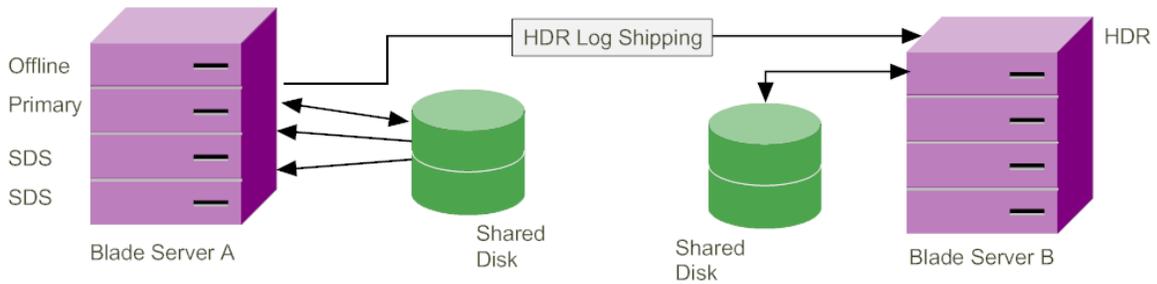
Figure 13. Transfer of the primary server to an SDS server with an additional HDR secondary server

Suppose, however, that the failure described in the previous illustration was not simply a blade within the blade server, but the entire blade server or the entire site. In this case you may have to fail over to the HDR secondary, and you could easily add additional SDS servers. Once the primary has been transferred to Blade Server B, then it becomes possible to start up SDS servers on Blade Server B as well, as shown in the following illustration.
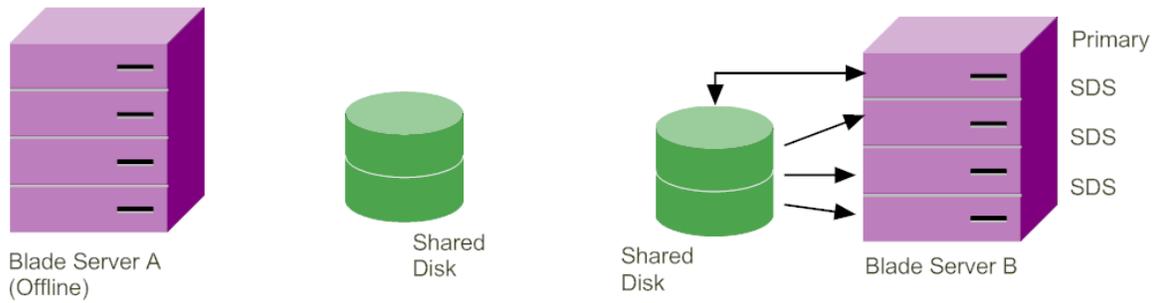
Figure 14. Transferring the primary to the HDR secondary and adding SDS servers

3.2   Using ER as Part of the Recoverable Group

While Enterprise Replication does not support a synchronous (SYNC) replication, it does provide the ability to support environments with multiple active servers. During a failover, Enterprise Replication is able to reconcile database differences with configurable conflict-resolution rules between the database servers. A configuration using Enterprise Replication is shown in the following illustration:
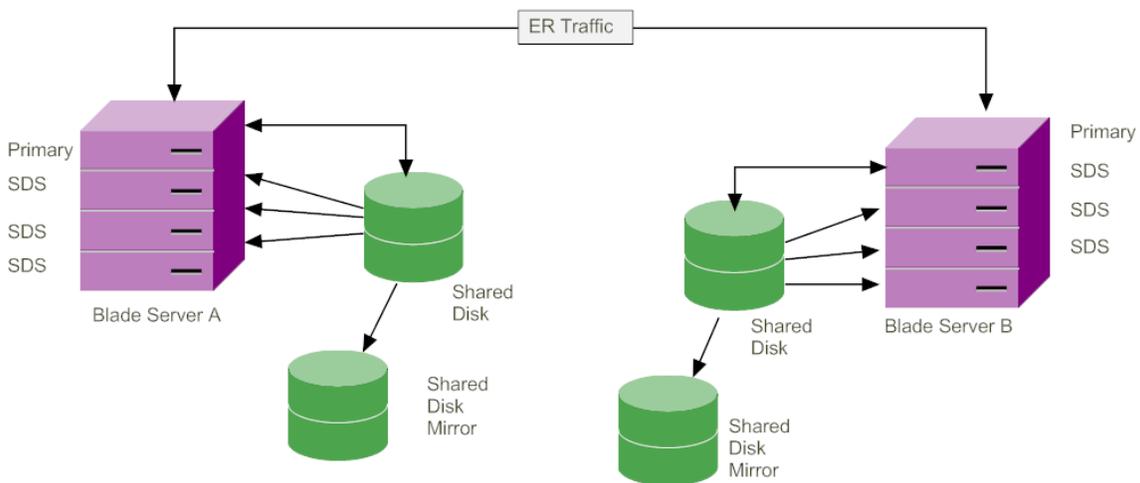
Figure 15. Using Enterprise Replication with SDS Servers

3. 3 Recoverable Groups Based on the HDR Secondary and the RSS Server

HDR secondary servers have provided high-availability for IDS customers worldwide. Over the past few years, customers have requested features such as:

- Additional HDR secondary servers
  Additional secondary servers can be useful for capacity relief where these servers can be added or subtracted to handle a changing workload, or for disaster recovery protection. One or more local servers help in case of hardware failure but if an entire facility were to experience a power outage or flood, then both the primary and secondary servers in that location would be lost at the same time. A secondary server at a remote site alleviates this risk and provides for business continuity, even in the event of a failure..

- Decreased impact on the primary from HDR replication

RSS servers and SDS servers directly address these requests. In addition to the extra workload relief that additional servers can provide, the RSS server operates with a fully-duplexed communication protocol. This means that the primary server does not have to wait for the RSS server to acknowledge the transmissions. This allows the RSS server to support a very high rate of data transfer even when there is high latency on the network. It also allows improved performance of the primary.
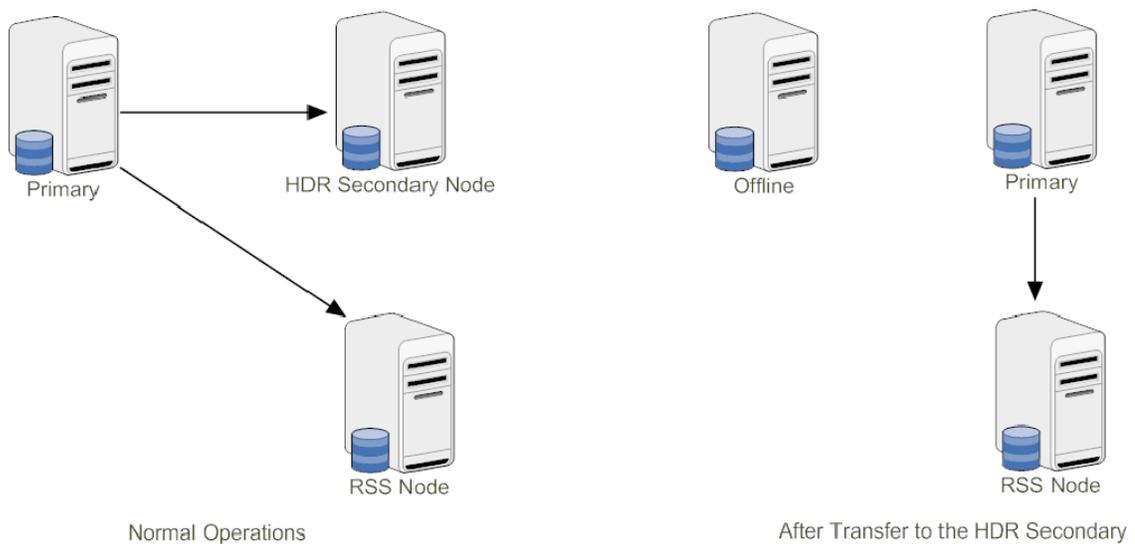
Figure 16. Using the HDR Secondary with the RSS

If it appears that the HDR secondary server will be off line for an extended period of time, the RSS server can be converted into the HDR secondary server. Then, when the original secondary server comes back on line, it can be converted to an RSS server as shown below.

Conversion of the RAS Node into an HDR Secondary   Bringing the Old Secondary Back Online
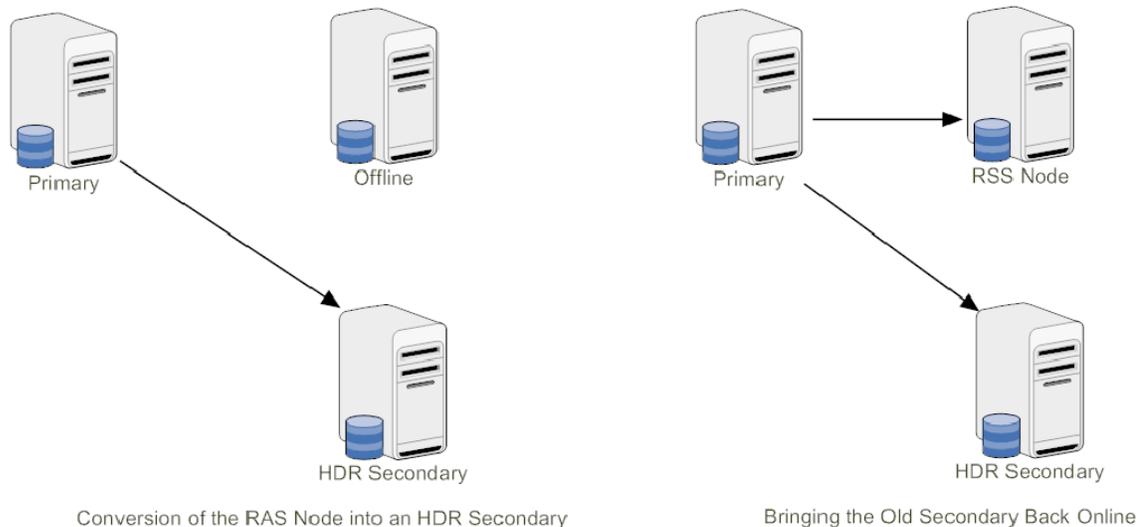
Figure 17. Conversion between the HDR Secondary and the RSS

When a system is configured with both an HDR secondary and one or more SDS servers, the HDR secondary provides backup in case the shared disk goes offline. If the system also has an RSS server then the RSS server provides additional availability for the HDR secondary. This configuration provides a true business continuity solution between HA and DR.

To provide continuous availability the availability configuration should be *layered*. The first layer provides availability solutions to deal with local failures. For example, this might include having a couple of blade servers attached to a single disk subsystem running SDS servers. Placing the SDS servers in several locations throughout your campus makes it possible to provide seamless failover in the event of a local outage.

You may wish to add a second layer to increase availability by including an alternate location with its own copy of the disks. To protect against a large regional disaster, you might consider configuring an HDR secondary server located some distance away. You may also want to make the remote system a blade server or some other multiple-server system. By providing this second layer, if a fail-over should occur and the remote HDR secondary became the primary, then it would be possible to easily start up SDS servers at the remote so that you still have a highly available configuration.

However, even a two-tiered approach might not be enough. A natural disaster in one region can trigger damage hundreds of miles away. To protect against this, consider

adding a third tier of protection, such as one or more RSS servers located thousands of miles away. This three-tier approach provides for additional redundancy that can eliminate the risk of an outage.
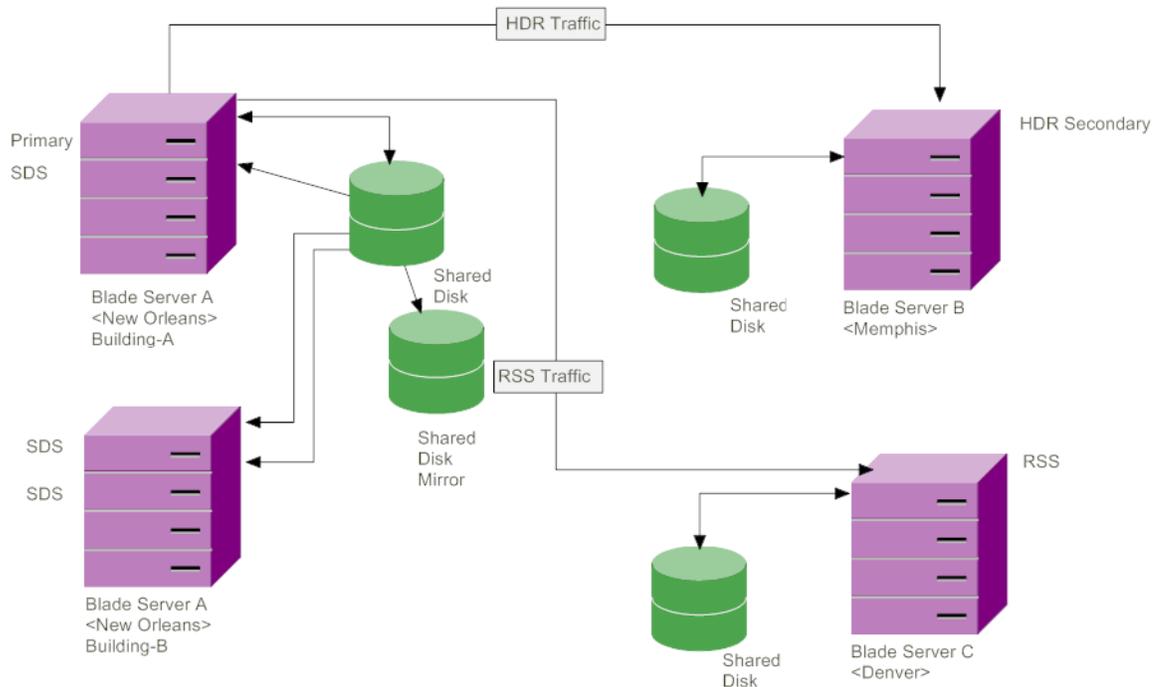
Figure 18. Preparing for multiple levels of failure

Referring to figure 18, suppose that a local outage occurred in Building-A on the New Orleans campus. Perhaps a pipe burst in the machine room causing water damage to the blade server and the primary copy of the shared disk subsystem. The role of the primary can easily be switched to one of the SDS servers running on the blade server in Building-B. This would cause all other secondary servers to automatically connect to the new primary server.
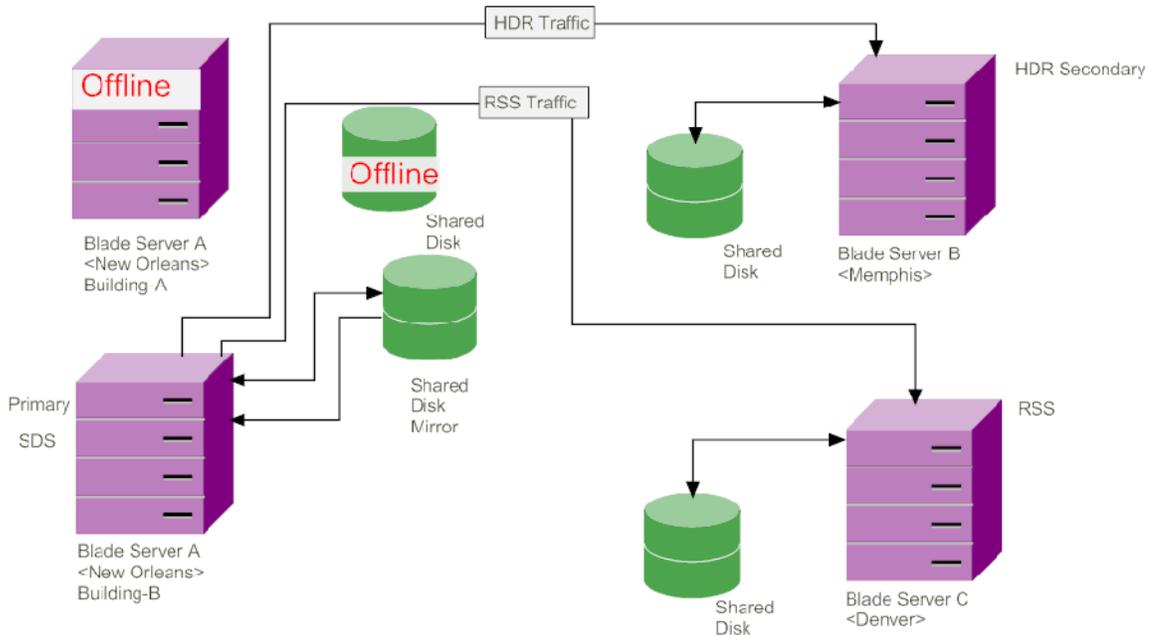
Figure 19. Failover after a Local Failure

Should there be a regional outage in New Orleans such that both Building-A and Building-B were lost, then Memphis becomes the primary server. In addition, you may also want to make Denver into an HDR secondary and possibly add additional SDS servers to the machine in Memphis. (Figure 20)
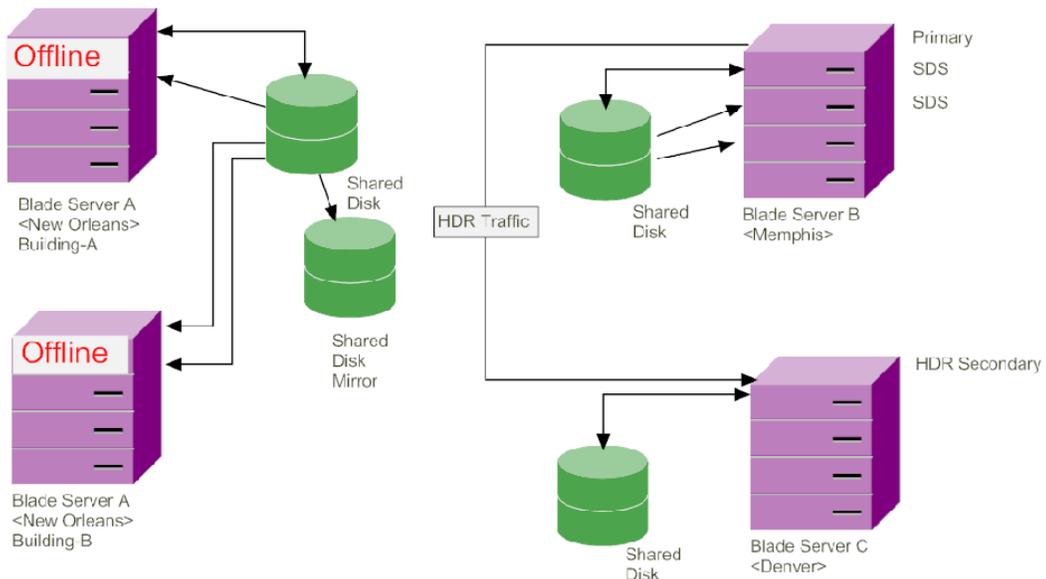
Figure 20.   Failover to Memphis

Finally, if the Memphis server is lost in addition to New Orleans, Denver could be converted to the primary server.

3. 3   Combining IDS Availability Options to Meet Different Needs

IDS provides many fundamental technologies for high availability. These technologies can be combined to suit a variety of business situations. Here are some examples of combinations that can be used to address different scenarios:

Table 1   Availability need and recommended solution

Application Requirements	Recommended Solution
You periodically want to increase reporting capacity	Use SDS or RSS servers. If the amount of data is very large and making multiple copies is difficult, SDS servers are more useful.
You are using SAN devices, which provide ample disk hardware availability, but are concerned about server failures	Use SDS servers and consider using a tiered approach with multiple types of secondary servers.
You are using SAN devices, which provide ample disk hardware mirroring, but also want a second set of servers that are able to be brought online if the main operation should be lost (and the limitations of mirrored disks are not a problem)	Consider using two blade servers running SDS servers at the two sites
You want to have a backup site some moderate distance away, but can not tolerate any loss of data during failover	Consider using two blade servers with SDS servers on the main blade servers and an HDR secondary in SYNC mode on the remote server
You want to have a highly available system in which no transaction is ever lost, but must also have a remote system on the other side of the world	Consider using an HDR secondary located nearby running SYNC mode and an RSS server on the other side of the world
You want to have a high availability solution, but because of the networks in your region, there is a large latency.	Consider using an RSS server
You want a backup site but you do not have any direct communication with the backup site	Consider using Continuous Log Restore with backup and recovery
You can tolerate a delay in the delivery of data as long as the data arrives eventually; however you need to have quick failover in any case	Consider using SDS servers with hardware disk mirroring in conjunction with ER
You need additional write processing power, can tolerate some delay in the delivery of those writes, need something highly available, and can partition the workload	Consider using ER for write processing with SDS servers for high availability

Summary

IDS provides many innovative features to support uninterrupted availability of data. While prior versions of IDS have provided high availability technologies, IDS 11 provides a full range of availability features that can be combined together according to business need, and can be quickly and easily reconfigured and expanded. The Multi-node Active Cluster for High availability (MACH 11) combinations and Continuous Log Restore greatly increase the flexibility and scalability of IDS while maintaining a very low cost of ownership.