Sun StorEdge™ Availability Suite Software – Compared With ORACLE Replication

A White Paper

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054 U.S.A.
650-960-1300

Part No.816-7182-10
June 2002, Revision A

Send comments about this document to: docfeedback@sun.com

Please
Recycle

Contents

Sun StorEdge Availability Suite Software – Compared With ORACLE Replication

This white paper focuses primarily on comparing ORACLE replication with the Sun StorEdge™ Availability Suite Software as used in a distributed database environment. ORACLE has positioned Oracle Data Guard for disaster recovery purposes. This paper does not go into details about Oracle Data Guard, but references it when necessary.

This white paper provides a general overview of Oracle replication and the Sun StorEdge Availability Suite Software. It does not discuss all the features of either product, which is beyond the scope of this paper.

ORACLE Distributed Database Architecture

A distributed database system helps enable applications to access data from local and remote databases. The databases can exist on the same host or on multiple hosts in multiple geographical locations. A set of databases in a distributed system can appear to applications as a single data source.

A distributed database system can be implemented as a client/server system and can comprise the following database types:

- All ORACLE databases (homogeneous)
- ORACLE databases and non-ORACLE databases (heterogeneous)

ORACLE Replication

The two terms, *distributed database system* and *database replication*, are related, but distinct. In a pure (that is, non-replicated) distributed database, the system manages a single copy of all data and supporting database objects. A distributed database system uses distributed transactions to access both local and remote data and to modify the global database in real-time. Database replication maintains multiple, distributed copies of the database, which helps improve local database performance and helps protect the availability of data to applications.

Replication is the process of copying and maintaining database objects in multiple databases that make up a *distributed database system*. Changes applied at one site are captured and stored locally before being forwarded and applied at each of the remote locations. Replication helps provide fast, local access to shared data, and helps protect the availability of the data to applications because alternate data access options exist. If one site becomes unavailable, you can continue to query or even to update the remaining locations. These and other benefits for applications are not possible within a pure distributed database environment.

Replication Environments

Some examples of the many designs for replication environments include the following:

- **Snapshot replication** - Master with read-only or updatable snapshot(s)
- **Advanced replication, single master** - Single master with snapshots that update the master
- **Advanced replication, multiple masters** - Multiple masters with no snapshots, synchronous or asynchronous replication modes
- **Hybrid** replication - Multiple masters with read-only or updatable snapshots, synchronous or asynchronous replication modes

Snapshot Replication

A snapshot is a full or partial replica of a master table at a single point in time. The snapshot is updated through individual batch updates, known as refreshes, from a single master site.

All snapshots provide the following benefits:

- Enable local access, which helps improve response times and availability for the main production site
- Helps offload queries from the main production site by enabling you to query the local snapshot instead
- Helps increase data security by enabling you to replicate only a selected subset of the target master table's data set

A snapshot can be a simple snapshot or a complex snapshot:

- **Simple snapshot** - A point-in-time copy of a table
- **Complex snapshot** - A point-in-time copy that is created, or refreshed, only if specified conditions and programmed select clauses are satisfied

Read-Only and Updatable Snapshots

A snapshot can provide read-only access to table data that originated from a master site. Read-only snapshots provide the following benefits:

- Help eliminate the possibility of conflicts because they cannot be updated
- Support complex snapshots
- Help enable you to avoid network access times
- Help provide continuous access to data when the network is down

Snapshots can be updatable. To ensure that a snapshot is consistent with its master table, you need to refresh, or update, the snapshot periodically. ORACLE provides the following three methods to refresh snapshots:

- **Fast refresh** - uses snapshot logs to update only the rows that have changed since the last refresh
- **Complete refresh** - updates the entire snapshot
- **Force refresh** - performs a fast refresh when possible, but when a fast refresh is not possible, performs a complete refresh.

Refresh Groups

When it is important for snapshots to be transactionally consistent with each other, you can organize them into refresh groups. By refreshing the refresh group, you can help ensure that the data in all of the snapshots in the refresh group correspond to the same transactionally consistent point in time. A snapshot in a refresh group still can be refreshed individually, but doing so nullifies the benefits of the refresh group because refreshing the snapshot individually does not refresh the other snapshots in the refresh group.

When creating a refresh group, you can configure the group so that ORACLE can automatically refresh the group's snapshots at scheduled intervals. Alternatively, you can omit scheduling information so that the refresh group must be refreshed manually. Manual refresh can be an ideal solution when the refresh is performed over a dial-up network connection.

Advanced Replication

Advanced replication supports symmetric, update-anywhere replication modes. All copies of the data can be updated and ultimately all sites converge on the same data.

ORACLE advanced replication supports the following types of replication environments:

- Single master replication

  Single master replication is a master site that supports one or more snapshot sites. All of the connected snapshots update only their master site. All data conflict detection and resolution occurs at the master site.

- Multimaster replication

  Multimaster replication, also known as peer-to-peer or n-way replication, enables multiple sites, where each site acts as an equal peer, to manage groups of replicated database objects in an update-anywhere model. Applications can update any replicated table at any site in a multimaster configuration. Updates made to an individual master site are propagated to all other participating master sites. ORACLE database servers operating as master sites in a multimaster environment automatically work to converge the data of all table replicas and to help ensure global transaction consistency and data integrity.

  When one or more sites update the same data, conflicts can occur. Conflict resolution is independently handled at each of the master sites. Multimaster replication provides complete replicas of each replicated table at each of the master sites.

Multimaster replication has two types: asynchronous and synchronous replication.

- **Asynchronous replication**, often referred to as store-and-forward replication, captures any local changes, stores them in a queue, propagates them, and applies them at remote sites at regular intervals. With this form of replication, there is a time lag (which you can vary) before all sites achieve data convergence. The changes are called deferred transactions.
- **Synchronous replication**, also known as real-time replication, applies any changes or executes any replicated procedures at all sites participating in the replication environment as part of a single transaction. If the transaction or procedure fails at any site, the entire transaction rolls back on all participating nodes. Synchronous replication helps ensure data consistency at all sites in real-time.

Procedural Replication

Batch processing applications can change large amounts of data within a single transaction. In such cases, typical row-level replication might load a network with many data changes. To avoid such problems, a batch processing application operating in a replication environment can use ORACLE's procedural replication. This process replicates the procedure (there might be an SQL statement along with the wrapper) to the remote sites, but it does not replicate the data modifications themselves. The remote site executes the procedure and applies changes to itself. Procedural replication can occur asynchronously or synchronously.

Benefits of Multimaster Replication

Multimaster replication has two major benefits: failover and load balancing.

Failover

Multimaster replication can be used to help protect the availability of a mission-critical database. For example, a multimaster replication environment can replicate all of the data in a database to establish a failover site should the primary site become unavailable due to system or network outages. In contrast with ORACLE's standby database feature, such a failover site also can serve as a fully functional database to help support application access when the primary site is concurrently operational, whereas a standby database can become fully functional only if the primary site is unavailable.

Load Balancing

Multimaster replication is useful for transaction-processing applications that require multiple points of access to database information for these purposes:

- Distributing a heavy application load
- Helping ensure continuous availability
- Providing more localized data access

Applications that have application load distribution requirements commonly include customer service-oriented applications.

Updatable Snapshots

You or your applications can insert, update, and delete rows of the target master table by performing these operations on the snapshot. An updatable snapshot can also contain only a subset of the data in the target master table.

Updatable snapshots have the following properties:

- Updatable snapshots are always based on a single table.
- Updatable snapshots can be incrementally (fast) refreshed.
- ORACLE propagates the changes made to an updatable snapshot to the snapshot's remote master table.
- If in a multimaster environment, ORACLE updates to the master table cascade to all other master sites.
- ORACLE can refresh an updatable snapshot as part of a refresh group in the same way it refreshes read-only snapshots.

Updatable snapshots provide the following benefits:

- Help enable users to query and update a local replicated data set even when disconnected from the master site.
- Require fewer resources than multimaster replication, while still supporting data updates.

Hybrid Replication

Hybrid replication is a combination of multimaster and snapshot replication. A single master site that supports one or more snapshot sites can also participate in a multimaster site environment, creating a hybrid replication environment. If the replication environment is a hybrid environment, the target master site propagates any of the snapshot updates to all other master sites in the multiple site replication environment.

ORACLE Replication Implementation Analysis and Limitations

This section describes other ORACLE implementation issues.

Replication Conflicts

Asynchronous multimaster and updatable snapshot replication environments must address the possibility of replication conflicts. These conflicts can occur when, for example, two transactions originating from different sites update the same row at nearly the same time. When data conflicts occur, you need a mechanism to ensure that the conflict is resolved in accordance with your business rules and to ensure that the data converges correctly at all sites.

In addition to logging any conflicts that might occur in your replicated environment, ORACLE replication offers a variety of built-in conflict resolution methods that help enable you to define a conflict resolution system for your database that resolves conflicts in accordance with your business rules. If you have a unique situation that ORACLE's built-in conflict resolution methods cannot resolve, you have the option of building and using your own conflict routines.

If a failure occurs at the primary site, recently committed transactions at the primary site might not have been asynchronously propagated to the failover site yet. These transactions appear to be lost. These lost transactions must be dealt with when the primary site is recovered.

High Availability and Survivability

The replication facility must be able to keep up with the transaction volume of the primary system.

For high availability, you must configure your system using one of the following methods. These methods are listed in order of increasing implementation difficulty.

- The failover site is used for read access only. That is, no updates are allowed at the failover site, even when the primary site fails.
- After a failure, the primary site is restored from the failover site using export/import, or through full backup.
- Full conflict resolution is employed for all data/transactions. This requires careful design and implementation. You must ensure proper resolution of conflicts that can occur when the primary site is restored, such as duplicate transactions.

- Provide your own special applications-level routines and procedures to deal with the inconsistencies that occur when the primary site is restored and when the queued transactions from the active failover system are propagated and applied to the primary site.

Replication Table

If possible, each replicated table should have a primary key. Where a primary key is not possible, each replicated table must have a set of columns that can be used as a unique identifier for each row of the table.

If any of the tables that you plan to use in your replication environment do not have a primary key or a set of unique columns, alter these tables accordingly. In addition, if you plan to create any primary key snapshots based on a master table, that master table must have a primary key.

Note – ORACLE does not support the replication of columns that use the LONG and LONG RAW datatypes. ORACLE simply omits columns containing these datatypes from replicated tables. You must convert LONG datatypes to large objects (LOB) in ORACLE8i.

ORACLE also does not support the replication of external or file-based LOBs (BFILEs). Attempts to configure tables containing columns of these datatypes as master tables return an error message.

Monitoring Tools

ORACLE provides a number of tools for monitoring purposes.

Replication Manager

Replication Manager is a GUI tool with which you can configure, schedule, and administer the replication environment. You can specify the objects, tables, indexes to replicate. You can also monitor system activity, resolve error conditions, and take action on any administrative issues.

Replication Catalog

You can check the status of jobs, broken jobs, database links, and replication errors using the replication catalog, which is internally maintained by ORACLE.

Oracle Data Guard

Oracle Data Guard is a tool for protecting an ORACLE database against disasters, human errors, and corruption. It is a powerful availability technology that allows a complete copy of a production database to be set up, administered, and maintained at a remote site. ORACLE Data Guard is included in the ORACLE database product. Some of the features of ORACLE Data Guard are:

- Integrates loosely-connected and geographically-dispersed production and standby databases into a robust, easily-managed disaster recovery solution. Oracle Data Guard provides the tools and infrastructures required to quickly install, instantiate, monitor, control, and maintain standby databases.
- Provides failover and switchover capabilities that allow easy role reversals between the primary and standby databases. This eliminates outages in database service due to planned maintenance as well as those resulting from a site disaster.
- Provides flexibility in data protection to suit specific business needs including synchronous mode with zero data loss or asynchronous mode with maximum performance.
- Safeguards against data corruptions and user errors. Using standby databases, primary-side physical corruptions, due to device failure, do not propagate through the redo logs that are transported to the standby database. Logical corruptions and user errors can also be prevented from propagating to the remote site.

Sun StorEdge Availability Suite Software

The Sun StorEdge Availability Suite Software includes two software packages: the point-in-time copy software and the remote mirror software.

Remote Mirror Software

The Sun StorEdge Availability Suite Software remote mirror software is a remote replication facility for the Solaris™ Operating Environment (Solaris OE). It is intended for use as part of a disaster recovery and business continuance plan to help provide redundant storage of critical information across physically separate sites.

The remote mirror software helps enable you to replicate disk volumes between physically-separate primary and secondary hosts in real time. To transport data, the remote mirror software uses any Sun network adapter that supports TCP/IP.

A remote mirror volume set consists of a primary volume on a local host and a secondary volume on a remote host. The volume set also includes a bitmap volume on each host to track write operations and differences between the volumes.

The remote mirror software enables you to group volume sets. You can assign specific volume sets to a group to perform replication on these volume sets and not on others you have configured. Grouping volume sets also helps guarantee write ordering in the asynchronous mode of replication across volumes. Write operations across the secondary volumes occur in the same order as the write operations to the primary volumes.

Types of Replication

The remote mirror software supports two types of replication: synchronous replication and asynchronous replication.

Synchronous Replication

A write operation is not confirmed complete until the remote volume has been updated. This forces the application to wait until the acknowledgement is received from the remote site.

Synchronous replication is more reliable and reduces the risk of data loss. However, there is an increase in response time, especially for large data sets or long distance replication.

Asynchronous Replication

A write operation is confirmed as complete before the remote volume has been updated. As soon as the data is written on the primary volume, the data is placed on the queue and the application is acknowledged for the completed I/O. The data is then transferred to the remote site and an acknowledgement is sent to the primary site.

Asynchronous replication provides fast response and has the least impact on the response time of the primary application. However, in-flight data loss can occur at the secondary site if a primary site or network failure occurs.

Synchronization (Resynchronization) Modes

The remote mirror software supports three synchronization modes: full forward synchronization, fast resynchronization, and full reverse resynchronization.

Full Forward Synchronization

Full forward synchronization is the process of initiating the full replication of the primary site to the remote site. When synchronization is complete (so that the primary and remote sites have the same data), the remote mirror software goes into replicating mode.

Fast Resynchronization

Fast resynchronization occurs when the delta data that has been changed but not yet replicated (for example: the writes that are captured while the volumes are in logging mode) are replicated across to the other side. The bitmaps on both sides are compared and the different bits' blocks are replicated.

Full Reverse Synchronization

During a fallback scenario, the data that has changed on the remote site after a failover can be synchronized back from the remote site to the primary site. This mode of synchronization does a full replication of the entire volume. When complete, the remote mirror software goes into normal replicating mode by replicating data from the primary site to the remote site.

Fast Reverse Synchronization

Fast reverse synchronization is the same as full reverse synchronization except that only the changed data is replicated back, not the full volume. The bitmaps are compared, and only those blocks that correspond to set bits at the remote site are replicated back to the primary site.

If for any reason, the replication must be stopped temporarily, the primary or secondary site could be forced to go into logging mode. The bitmaps keep track of any changes happening at the active site. Later, the data can be replicated to the other site using fast syncs.

Advanced Features of Remote Mirror Software

The remote mirror software supports several advanced features: grouping of volume sets, one-to-many replication, and multi-hops.

Grouping

A group is a collection of remote mirror software volume sets that have the same group name, primary and secondary interfaces, and mirroring mode. Mixed groups, groups in which mirroring modes are asynchronous for one set and synchronous for another set, are not allowed.

You can group volume sets together so that the remote mirror software preserves write ordering across the volumes in asynchronous mode replication. For example, you can group the different ORACLE data files and easily maintain the group for various remote mirror options.

You can also control remote mirror software volume sets simultaneously by grouping them. This feature is essential in installations for which you must maintain consistent contents of a group of volumes.

One-to-Many Replication

You can replicate data from one primary volume to many secondary volumes that reside on one or more hosts. When you perform a forward resynchronization, you can synchronize one volume set or all volume sets by issuing a separate command for each set. You can also update the primary volume from a specific secondary volume.

You can replicate one primary site volume to more than one site using different modes of synchronization. If the network on one site fails, replication is forced into logging mode only for that site. Replication continues for other sites.

Multi Hops

A secondary site can become a primary site and it can replicate to a third site as its secondary site.

Using Remote Mirror Software With ORACLE

When you use the remote mirror software with ORACLE replication, you can either replicate only the redo logs or you can replicate the full database.

Replicate Only Redo Logs

Consider an environment where the production database runs on the primary site and a standby database is maintained at a remote secondary site. Archive logs are shipped using regular ORACLE procedure, and the online logs are replicated using the remote mirror software in real time.

If the primary site or network is down, the standby database can be recovered to the latest transaction, depending on whether the remote mirror software is in synchronous mode or asynchronous mode. If the software is in synchronous mode, the latest transaction can also be recovered; if the software is not in synchronous mode, there might be a slight transaction loss (depending on network latency and bandwidth).

Replicate the Full Database

The complete database can be replicated in a combination of synchronous and asynchronous modes. Replicate online logs in synchronous mode and replicate the grouped data files in asynchronous mode.

If the primary site or network is down, the instance at the remote site can be started and users can access the remote site.

When the problem is fixed, the data from the remote site can be reverse-synchronized to the primary site, and the database can then be restarted at the primary site.

You can also reverse the scenario by switching the roles of the primary and the secondary sites.

Point-in-Time Copy Software

The Sun StorEdge Availability Suite Software point-in-time copy software is a snapshot volume copy facility for the Solaris OE. With the point-in-time copy software, you create shadow volume sets, which consist of a master volume, a shadow volume that contains a point-in-time copy of the master, and a bitmap volume to track changes in the volumes. After the shadow volume is established, you can read from and write to both this shadow volume and the master volume.

The point-in-time copy software enables you to quickly update the shadow volume from the master volume or to restore the master volume from the shadow volume. The software also supports fast resynchronization, which enables you to create a new point-in-time volume copy by updating the specified volume with only the changed data.

The point-in-time copy software can be used with ORACLE databases to create point-in-time copies of the database that are available for use almost immediately.

The volumes that need to have a snapshot taken are defined initially. Then the tablespaces are put into hot-backup mode and the snapshot command is issued. As soon as the command prompt returns, the tablespaces can be taken out of hot-backup mode.

Snapshot Modes

The point-in-time copy software supports several volume set types that enable different snapshot modes: a dependent shadow volume set, an independent shadow volume set, and a compact dependent shadow volume set.

Dependent Shadow Volume Set

Only changed data is written to the shadow volume. The snapshot is nearly instantaneous because no copy is performed upon creation of the shadow volume set. Attempting to read unchanged data from the shadow volume results in the read coming from the master volume. The shadow volume size is equal to or greater than the master volume. A disadvantage is that if the master volume is not available or is corrupted, the shadow copy also become useless.

Independent Shadow Volume Set

The full master volume is replicated to the shadow volume upon creation of the shadow volume set. The shadow volume can be read while the copy is proceeding because the set behaves as a dependent shadow volume until the copy is complete.

One advantage is that the shadow volume, after the copy is complete, is independent of the master and a failure of master volume doesn't affect the availability of the shadow.

Compact Dependent With Overflow Volumes

If you know that the data changes to the master volume are going to be minimal, then you do not need to allocate the shadow volume to be the same size as the master volume. In such cases, you can create a compact dependent shadow volume set, in which the shadow volume is smaller than the master volume. Using compact dependent shadow volumes is especially beneficial when the master volume is very large and the expected change is small. If the dependent copy is full, you can attach another volume to the compact shadow volumes, which acts as an overflow volume. You can attach an overflow volume to the compact dependent shadow volume set to protect against an unexpected volume of changes to the master filling up the compact shadow volume.

Point-in-Time Copy Software Operations

The point-in-time copy software supports a number of methods for maintaining volume copies and for providing access to these volumes.

Full Update

The master volume is fully copied on the shadow volume in an independent shadow volume set.

Fast Update

To quickly update the changes on the master volumes into the shadow volumes. In an independent shadow volume set, the changes are physically copied. In a dependent shadow volume set, the data is cleared from the shadow and the bitmaps point to the master data.

Reverse Update

Reverse synchronize the data from the shadow volume to the master volume. Reverse update is useful when a master volume is corrupted and needs to be restored. Instead of reading from tape, the data is quickly restored from the shadow volume and the master volume is recovered.

Export/Import/Join

This feature of the point-in-time copy software provides another way to take data from the shadow volume set. The shadow volume is exported from the shadow volume set so that a second host can import it and use it in another instance. Later, the shadow volume can be released from the second host and the shadow volume can be joined to its original shadow volume set. The shadow volume must reside on a multiported device.

Point-in-Time Copy Software Features

The point-in-time copy software supports two additional features: grouping of volume sets and multiple shadow volumes of the same master volume.

Grouping

You can organize volume sets in groups for ease of administration. Grouping enables atomic execution of commands across all group members. Grouping helps assure consistent point-in-time copies across all members of a group.

Multiple Shadows

You can enable more than one shadow volume set using a single master volume, which gives the master volume multiple shadow volumes. You can use each of these shadow volumes independently for software evaluation or data analysis.

Implementing the Sun StorEdge Availability Suite Software

ORACLE database availability is even higher when the database is implemented with the Sun StorEdge Availability Suite Software.

The following sections give the various options that are available.

Method 1: Snapshot / Replicate / Snapshot

1. Enable a point-in-time copy independent shadow volume set at the remote mirror primary site, and keep the set intact.
2. With the remote mirror software, replicate the shadow volume to the remote site. A full synchronization is performed initially, but it can be a fast synchronization later.
3. When the remote mirror replication is complete, put the primary and secondary remote mirror volumes into logging mode.
4. At the remote site, enable an independent point-in-time copy shadow volume set with the remote mirror secondary volume as the master volume.
5. You now have three complete sets of data, which removes the single point of failure.
6. The data at the remote site can be used for backup, for test, or for data warehouse processing.

Method 2: Replicate / Snapshot

1. With the remote mirror software, replicate the primary volume to the secondary volume at the remote site.
2. Put the remote mirror volumes in logging mode.
3. Enable a point-in-time copy independent shadow volume set at the remote site with the remote mirror secondary volume as the master volume.
4. Put the remote mirror primary and secondary volumes into replicating mode again.
5. Use the point-in-time copy shadow volume at the remote site for backup, for test, or for data warehouse processing.

Method 3: Snapshot / Replicate

1. Enable a point-in-time copy independent shadow volume set at the remote mirror primary site and keep the set intact.
2. With the remote mirror software, replicate the shadow volume to the remote site. A full synchronization is performed initially, but it can be a fast synchronization later.
3. When remote mirror replication is complete, put the primary and secondary remote mirror volumes into logging mode.
4. The data at the remote site can be used for backup, for test, or for data warehouse processing.

Comparing ORACLE Replication With the Sun StorEdge Availability Suite Software

Multimaster replication is analogous to remote mirror software and snapshot sites are analogous to point-in-time copy software.

- The Sun StorEdge Availability Suite Software is meant for disaster recovery whereas ORACLE replication is primarily implemented as an application requirement such as offloading data and distributed database.
- The Sun StorEdge Availability Suite Software works at the volume level whereas ORACLE replication works at the object or row level. ORACLE logically replicates columns or rows that have changed, but remote mirror software replicates blocks that have changed on the replicating volume.
- The Sun StorEdge Availability Suite Software helps provide high availability for any database, whereas ORACLE replication software is a bundled product that works only with ORACLE databases.
- Both the remote mirror software and ORACLE replication handle synchronous and asynchronous replication.
- The remote mirror software is straightforward and easy to maintain, while the ORACLE replication environment can be complex.

Advantages of ORACLE Replication

- The product is bundled with the ORACLE Standard and Enterprise editions at no extra cost.
- Works in a heterogeneous environment with multiple operating systems and types of hardware. For example, a sales person can download data into a laptop that is running a different operating system when visiting a customer. Later, the master data can be updated by replicating the data from the laptop to the server.
- Makes many-to-many real-time updates possible.
- Works at the table and row level rather than at the data file level.
- All the instances in the distributed environment can access the data and the data is propagated to other sites.
- No need to put the tablespace into backup mode while taking snapshots.
- Can selectively pick up the tables to replicate or to take snapshot for different sites.

- Production environments can see an improvement in performance because the local data snapshot is available for other snapshot sites, which helps reduce the load on the production database.
- The target database is available for query.
- Since data is logically replicated, data corruption is not replicated from the primary site.

Disadvantages of ORACLE Replication

- Does not replicate data types such as Bfile, Long, and LongRaw. Also, to replicate schema objects with user-defined types, the user-defined types must exist on all replication sites and must be exactly the same at all replication sites.
- Changes in data structures on master sites need to be coordinated with the other master and snapshot sites, especially in a distributed environment.
- Increased risk of replication failure with large refresh volumes because replication cannot occur until the transaction is complete.
- Synchronous multimaster replication can be implemented only when it is absolutely certain that the network is very stable.
- ORACLE replication is often implemented in a distributed database environment where each database is part of the complete database and is not used as a standby for any other database. In this environment, it is not a disaster recovery tool.
- When ORACLE replication is used for disaster recovery, the multimaster environment is implemented at a complex level, which can be difficult to maintain. See “Oracle Data Guard” on page 10.
- Failover sites in a replicated environment are often referred to as read-only sites, because updatable failover sites can be difficult to implement.
- Cannot restore the complete database from one site to another.
- Does not replicate standalone procedures.
- Need to maintain privileges across sites.
- If used with the `nologging` and `direct` options, the transactions and data loading is not logged and might not be replicated.

Advantages of the Sun StorEdge Availability Suite Software

- Can be used in the development environment to replicate source code, binaries, and libraries for faster restoration.
- Extremely simple to implement and maintain.

- Not as complex as ORACLE replication; easier to understand.
- Can replicate almost any data, not just ORACLE databases.
- Because it is mainly used as a disaster recovery tool and to reduce the load from production databases while doing backups, the software is focused on availability and reliability.
- Doesn't need to know the structural changes that have happened at the database. It replicates the data at the volume level.
- Failover and fallback is straightforward, easy, and fast.
- Requires minimum maintenance for the shadow volume sets in operation.
- The ability to do fast syncs, forward and reverse, simplifies data file restoration, which helps minimize the mean-time-to-recover (MTTR).
- The replicated snapshot data is used for backup to tape, which lessens the load on the production database. The database must be in hot-backup mode for only a very short time compared to the traditional backup done on production databases.
- Can maintain the replicated data at more than one location, which helps increase the availability of the data.
- One-to-many and multi-hop capabilities with the remote mirror software and one-to-many snapshots are possible.

Disadvantages of the Sun StorEdge Availability Suite Software

- Runs only in the Solaris OE.
- Replication happens at the volume level, therefore single file restoration on a mounted volume is not possible: it will restore all the files on the volume.
- With remote mirror software, while the replication of the full database is going on, the remote site's data cannot be used. It has to be in logging mode before it can be used. However, by taking a point-in-time snapshot on either the primary system or the secondary system, you can access the shadow volume’s data while the replication is proceeding.
- If UFS or VxFS is used and replication is in progress, the remote volume or shadow volume should be in an unmounted state. With a point-in-time copy, after the copy is complete, the shadow can be accessed.
- Many-to-many online replication or snapshot is not possible.
- Since there is an additional layer of software, there is an additional performance overhead.

Conclusions

- If you are considering a serious disaster recovery process, the Sun StorEdge Availability Suite Software can be a better choice than ORACLE replication.
- If ORACLE replication is used for distributed processing, then the Sun StorEdge Availability Suite Software can complement the environment at one of the master sites, but cannot replace ORACLE replication.
- If the customer implemented basic ORACLE replication to snapshot just a few tables, then the Sun StorEdge Availability Suite Software would not be of great use.
- Based on the various advantages and disadvantages of both products, and depending on the particular situation, the customer can:
  - Implement the Sun StorEdge Availability Suite Software for database and development environment to provide business continuance in the event of a site failure.
  - Implement ORACLE replication for distributed databases.
  - Implement both the Sun StorEdge Availability Suite Software and ORACLE replication in a complementary configuration.
  - Implement Oracle Data Guard and the Sun StorEdge Availability Suite Software, if a standby is going to be maintained at the secondary site.

References

- - ORACLE 8.1.6 server documentation

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054 USA
650-960-1300 Fax 650-969-9131

For U.S. sales office locations, call: 800-821-4643

In other countries, call:
Corporate headquarters: 650-960-1300
Intercontinental sales: 650-688-9000

Contact: Sridhar Ranganathan
Network Storage
Sun Microsystems, Inc.

Comments: Please write to:
sridhar.ranganathan@sun.com