Building a 100% Reliable Backup & Recovery for Oracle Apps

Dan Hotka – Quest Software

Introduction

Over the last few years, corporations have invested billions of dollars to integrate the automations of core business systems into large Enterprise Resource Planning (ERP) applications. Increasing reliance on the availability of ERP environments and the advent of "around the world, around the clock" business transactions via e-commerce exposes organizations to a great risk. Losing access to the ERP system or the e-commerce application for an extended period of time may cause the entire business to collapse. However, this is only one part of the picture. Most people in decision-making positions are totally dependent on the application to achieve their productivity goals. This means that periods of poor application response time directly impact the organization's productivity. This paper proposes measures to reduce the risk of downtime while increasing the day-to-day application response time by combining multiple state-of-the-art technologies.

Definition of 24x7

When evaluating the need for high availability, you must first define what 24x7 means to your organization. Here is a list of questions that will help you determine your 24x7 needs:

1. Can you assign a dollar amount for application downtime?

   Some applications may be internally or externally regulated with predefined penalties for periods of lack of service. For other applications, you must quantify the direct and indirect costs of downtime.

2. Do all application components/processes share the same level of importance?

   Some applications may have different requirements for different parts of the application. For example, data entry may be available only between 8:00 a.m. and 5:00 p.m. A particular batch job may need to be completed before 7:00 a.m. Such requirements may mean that although the criticality of the entire application may not warrant the high cost associated with higher availability, some investment should be made to combat certain failures, especially during a certain time frame. For example, with 2,400 data entry clerks working from 8:00 to 5:00, it would be important that during this time the application does not experience more than two minutes of downtime. However, after 6:00 p.m., the criticality of this data entry application diminishes until the next morning.

3. Is the availability of the application more important than data consistency?

   In financial applications, data integrity is paramount. Under no circumstances should a committed transaction be lost, even if it means more downtime for recovery. In many order entry systems, it is more important that the application remains available at all times even if a few orders are lost. This is especially true for e-commerce style applications. For many on-line shopping applications, it is most important to have the system accept new orders at all times. If the system goes down, another system must take over even if it means that several transactions may be lost. The cost of downtime due to recovery is greater than the cost of lost transactions.

4. When can you do application and database upgrades?

   Is there a natural time during which the availability of the system is not critical? Is it long enough to accommodate an upgrade? How often is that quiet period of time — nightly? Every weekend? Once a month? On national holidays? A different solution may be required for systems that may be maintained during the night versus systems that do not have such natural quiet time. If your systems do not have such quiet times, there should be some administrative time allocated.

Performance Impact on Availability

Downtime is not the only concern when dealing with high availability. Most large applications encounter diverse processing needs. On one hand, the online users require good response time from many short transactions. On the other hand, large batch jobs (e.g., reports and complex extracts) expect high throughput of a handful of very large transactions. These conflicting needs cause response time to fluctuate, decreasing the reliability and availability of the application. This is especially a concern with applications that provide services directly to end users and consumers, such as an e-commerce application.

Redundancy is the Key to Availability

The logical solution for increased availability is to maintain the data in more than one place. Coincidentally, one of the best solutions for improving the application response time is to separate the batch reporting and extract processing from the OLTP processing. This solution also requires a redundant database. This means that owners of large critical applications are seeking a solution that provides both high availability and the ability to offload noncritical processing from the main system.

The criteria for a comprehensive high availability and high performance solution include:

- Minimal impact on the primary system — the solution should have limited impact on the availability and the performance of the primary system.

- There should be a full copy of the primary database, which should be accessible even when there is no emergency — mainly for reporting and extracts.

- The copy of the database should be an up-to-date image of the primary database.

- When a disaster occurs, the copy should be capable of becoming the primary database (fail-over).

- After the disaster, the solution should take into account the switch back to the primary system.

- It is desirable to be able to modify some aspects of the copy database to accommodate the different processing on it. For example, special indexes should be built to support the reporting needs. These modifications should be reversible when a fail-over occurs.

- It is desirable that the copy will not require its own database administration in addition to the administration of the primary system.

- It is desirable that there will be redundancy in CPU as well as in the database.

- It is desirable that the secondary system will be located at a remote site.

- A fail-over to the secondary database should be very fast, and no data should be lost during the process.

Range of Common Solutions

There is a wide range of solutions to the high availability problem. The most common methods are:

- **Local disk mirroring and/or RAID** – This solution does provide protection against many disk-related failures, but the mirror is usually not accessible under normal circumstances. Once broken, the mirror becomes stale relative to the copy that is still operational. To resync (or re-silver), many disk mirroring solutions perform a complete copy of the data from the operational copy to the stale copy. If the database is large, this process can take a very long time. Other disk mirroring techniques such as those provided by EMC and Veritas provide for a delta refresh, which is much faster. Local disk mirroring does not provide a resolution for a local disaster. It also does not provide protection against physical block corruption by Oracle or accidental loss of data due to a DBA error (such as dropping or truncating a production table).

- **Oracle standby database** – This solution provides some protection against a catastrophe that makes the primary database unavailable. However, it is notoriously unreliable. Oracle's standby database has other shortcomings. The copy is only as current as to when the last log was applied. Once the database is opened and modified, a complete image is required to reconstruct the standby database. Additionally, there is some administration required for the standby database as the structure of the source database changes, such as adding data files or auto-extending tablespaces.

- **Local clustering** – Local clustering is a hardware solution that enables multiple computers to share a set of disks. Applications on these computers are written such that they can freely migrate between the machines in the clusters using a technology known as "floating IP addresses." Unfortunately, the Oracle database relies on persistent memory structures, so when a switch happens, the database has to be brought down and restarted.

  This solution provides good protection against most common failures. However, since there is only one copy of the database, there should still be consideration for protection of the disks. Moreover, since there is only one copy of the database, any physical block corruption or accidental dropping of a database object will cause the application to fail. Finally, with a local cluster, there are no provisions for performance improvement by any load sharing.

- **Remote disk mirroring** – Remote disk mirroring provides the protection of disk mirroring, but to a remote location. Remote disk mirroring comes in two flavors: synchronous and asynchronous. With asynchronous mirroring, the primary system does not wait for the data to be "committed" to the remote disk. With synchronous mirroring, the application waits for the data to be committed to both the local and the remote disk. When the application is an Oracle database, asynchronous mirroring may cause a situation that can cause the mirrored database to be structurally corrupted. This means that when a switchover is required, the DBA will not be able to open the remote database. For this reason, most remote mirroring implementations use the synchronous method. However, with this method, a wide bandwidth is needed between the source and destination. Otherwise, the network will slow the primary system. Most sites use the remote disk mirroring from EMC with one or more T3 lines.

- **Replication** – Replication provides the ability to have a live remote database that can be used both to reduce the workload of the primary system and for fail-over when a disaster happens. However, Oracle replication is resource-intensive, which causes it to have a substantial impact to the primary system. Moreover, because of the many limitations imposed by Oracle replication, it cannot successfully replicate many of today's applications. In particular, you cannot replicate large ERP sites. SharePlex® for Oracle from Quest Software overcomes these problems. SharePlex for Oracle is a comprehensive and efficient solution that can support replication for most ERP sites. The live database on the remote site does require database administration, and application of patches to the application is not straightforward.

- **Local clustering with Oracle Parallel Server** – Oracle Parallel Server (OPS) offers another alternative for high availability systems. Using this facility, many instances of Oracle running on different hardware can access the same database on shared disks. This does permit the hardware that would be allocated for a standby system to be actively used in production. Concurrent access to data by the different instances is managed by the Distributed Lock Manager (DLM), an application that assures that an instance always accesses a consistent block from the database. The DLM causes the instance holding a dirty data block to flush it to disk, so the reading instance can access a clean copy. The difficulty in using OPS for highly available solutions is that the application needs to be designed so that transferring blocks between instances (pinging) is minimized. If not, application performance can be severely degraded. Also, with OPS, there is only one copy of the database that is not protected from disk failures or corruption.

The Integrated High Availability and High Performance Solution

From the brief description of the high availability options shown above, it is clear that there is no one solution that can support all the requirements put forth above. To solve the problems of both high availability and the need for high performance several vendors had to combine their forces. These vendors are EMC, HP and Quest Software. The solution is to provide a local cluster of machines with local and remote disk mirroring which provides physical protection against hardware failure. An on-line replica of the database is available for off loading of reports, and it provides logical redundancy. The on-line replica provides protection against logical blocks corruption, accidental database object drops and security breaches while protecting the performance of the main production environment by offloading non-critical processing.

Implementing the Solution

HP provides the clustering solution with **HP MC/ServiceGourd**. The cluster automatically detects the loss of an application or any critical part of it. It then executes a script that switches over the application to a secondary machine.

EMC provides the disaster tolerant disk solution with the **Symmetrix**™ disk drives. These disks operate on the production system. A secondary copy of the disk is maintained via remote disk mirroring in a disaster recovery site with **SRDF**™.

Quest Software provides the logical replication solution with **SharePlex® Overdrive**. The replication is usually done to a third mirror that is created via EMC's **TimeFinder**™ software. The unique integration between the logical and physical replication/mirroring solutions enables a physical synchronization of the logical replica to be done automatically at any time with no impact on the source system.

Introducing SharePlex Overdrive

Quest Software's SharePlex Overdrive is designed to provide a continuous replication solution to databases that support large applications. SharePlex is optimized to minimize the impact on the source system, the source instance and the network. SharePlex Overdrive includes:

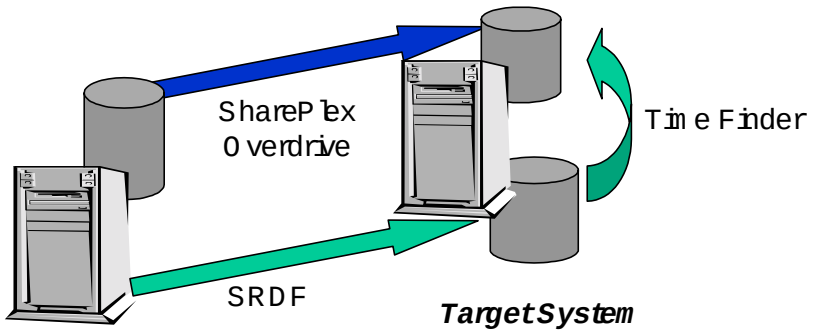
- **SharePlex for Oracle** – provides high-performance Oracle replication that maintains a continually updated, accessible target instance

- **SharePlex Reconcile Post** – enables SharePlex Overdrive to continue replication appropriately after a fast EMC refresh of the underline disks. This enables quick synchronization between the replica and the source without impacting the source system itself.

- **SharePlex Monitoring** – providing unattended monitoring of SharePlex services using a component of Quest Software's I/Watch™.

Providing the best in disaster recovery with useful, up-to-date secondary instances for distributed processing, EMC and Quest Software are delivering the best of both worlds.

The Combined Solution



EMC's SRDF performs disk-level replication of the entire environment from the source system to the target system. In synchronous mode, no change is made to the source system that isn't also made to the target, guaranteeing that the target system's second mirror is an exact replica of the production server. If the production server fails, a secondary system is available with the same data, and the disk drives need only be mounted in order for business to continue despite the outage of the primary server.

EMC's TimeFinder creates the Oracle database for the "third mirror," quickly generating the copy that SharePlex Overdrive then maintains by replicating changes to selected tables and sequences from the source system to the third mirror.

In addition to greatly reducing the time required to create the initial image for the third mirror, TimeFinder is used to refresh the Oracle database periodically or when the database structure changes. This ability to use EMC's fast refresh capability of TimeFinder eliminates the need for database administration on the replica. However, since the replica is a live independent database, many sites go further and modify the replica to facilitate the processing that take place on the replicated image. In particular, many sites create new indexes on the replica to support the application reporting needs. These indexes are reestablished after every TimeFinder sync.

Benefits of the Proposed Solution

In the standard SharePlex Overdrive scenario, a customer has Symmetrix disk drives on the source and target systems, with EMC's SRDF between these disks. A third mirror is created via EMC's TimeFinder, which is maintained as an accessible, up-to-date reporting instance by Quest Software's SharePlex Overdrive. This combination affords many benefits:

- **Easy initialization for 24x7 shops:** SharePlex Overdrive requires two matching copies of data with which to start replication. SRDF and TimeFinder, in conjunction with SharePlex Overdrive, make this requirement easy to fulfill because SRDF and TimeFinder can create an initial replica instance, and SharePlex's integration module for EMC will reconcile that instance with the information contained within SharePlex's queue files.

- **Reduced disaster recovery time**: Since SharePlex Overdrive maintains an available standby instance, if a disaster occurs on the source system, this standby instance can be used until the application can be restarted on the disaster recovery copy maintained by SRDF. This can reduce the application downtime significantly.

- **Provides more disaster recovery protection**: SRDF, like every other mirroring solution, is

prone to block corruption errors and to human errors such as accidentally dropping a production table. Since the replica maintained by SharePlex is a logical replica of the database, it provides a protection against both block corruption (known as Oracle Error 600) and accidental erroneous DDL.

- **Fast, easy migration without downtime**: SharePlex Overdrive can replicate between different versions of Oracle. When you need to upgrade an Oracle version, you can perform the upgrade on a secondary system. SharePlex for Oracle will keep the data current on the upgraded database. Once the upgrade has been fully tested, you can use EMC's fast refresh capability to upgrade your production system to the new version. SharePlex for Oracle minimizes downtime to the production system.

- **Flexible configurations with WAN support**: SharePlex Overdrive can replicate between instances on the same system, to instances on a local area network, or to remote instances through a wide area network. Additionally, through a cascading scenario, SharePlex Overdrive can replicate from one system to several and from those onto further systems. With this type of configuration, a global company could replicate from New York directly to Boston, D.C., Atlanta, and London, and then have the London office replicate to Madrid, Paris, Rome, Brussels, and Berlin, limiting the traffic across the ocean while keeping all remote offices up to date.

****

Dan Hotka is a Director of Database Field Operations for Quest Software. He has over 21 years in the computer industry and over 16 years experience with Oracle products. He is an acknowledged Oracle expert with Oracle experience dating back to the Oracle V4.0 days. He has co-authored the popular books Oracle Unleashed, Oracle8 Server Unleashed, Oracle Development Unleashed by SAMS and Special Edition using Oracle8/8i by Que, is frequently published in trade journals, and regularly speaks at Oracle conferences and user groups around the world. Dan can be reached at dhotka@earthlink.net or dhotka@quest.com .