

Oracle Database 11g High Availability

An Oracle White Paper
June 2007

NOTE:

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Oracle Database 11g High Availability

Introduction	3
Causes of Downtime	3
Computer Failure Protection.....	4
Real Application Clusters.....	5
Bounding Database Crash Recovery Time	6
Data Failure Protection	6
Storage Failure Protection	7
ASM Block Repair.....	7
Rolling Upgrades of ASM.....	8
Site Failure Protection.....	8
Data Guard.....	8
Human Error Protection	11
Guarding Against Human Errors	12
Oracle Flashback Technology	12
Data Corruption Protection	15
Oracle Hardware Assisted Resilient Data (HARD)	16
Backup and Recovery	16
Planned Downtime Protection.....	18
Online System Reconfiguration.....	19
Online Patching and Upgrades.....	19
Online Data and Schema Reorganization	22
Maximum Availability Architecture – Best Practices	24
Conclusion.....	24

The increasing demand on IT within the enterprise has established a critical relationship between business success and the availability of the IT infrastructure.

INTRODUCTION

Enterprises leverage Information Technology (IT) to garner competitive advantage, reduce operating costs, enhance communication with customers, and increase management visibility into core business processes. As the use of IT and IT enabled Services (ITeS) become more and more pervasive in all aspects of business operations, modern enterprises are highly dependent on their IT infrastructure to be successful. Unavailability of a critical application or data may have a significant cost to enterprises in terms of lost productivity and revenue, dissatisfied customers, and tarnished corporate image. A highly available IT infrastructure is therefore, a critical success factor for businesses in today's fast moving and "always on" economy.

The traditional approach to building high availability infrastructure requires widespread use of redundant and idle hardware and software resources supplied by disparate vendors. Such an approach is not only very expensive to implement, it also falls short of meeting user's service level expectation due to loose integration of components, technological limitations, and administrative complexities. Responding to these challenges, Oracle has been working hard to provide customers with a comprehensive set of industry leading high availability technologies that are pre-integrated and can be implemented at a minimal cost.

In this paper, we will review the common causes of application downtime and discuss how technologies available in the Oracle Database can help avoid costly downtime and enable rapid recovery from unavoidable failures. We will also highlight some of the new technologies introduced in Oracle Database 11g that enable businesses to make their IT infrastructure even more robust and fault tolerant, maximize their return on investment on High Availability infrastructure, and provide better quality of service to users.

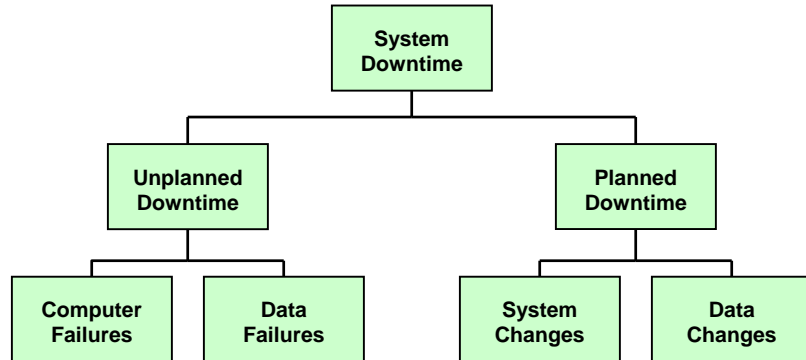
Causes of Downtime

When architecting a highly available IT infrastructure, it is important to first understand the various causes of application outages. As depicted in Figure 1 below, downtime can primarily be categorized as unplanned and planned. Unplanned outages are generally caused by computer failures as well any other

It is critical to understand the various causes of application downtime in order to architect an effective high availability architecture.

failures that may cause the data to be unavailable (e.g. storage corruption, site failure, etc.). System maintenance activities such as hardware, software, application, and/or data changes are typical causes of planned downtime.

Figure 1: Causes of Downtime

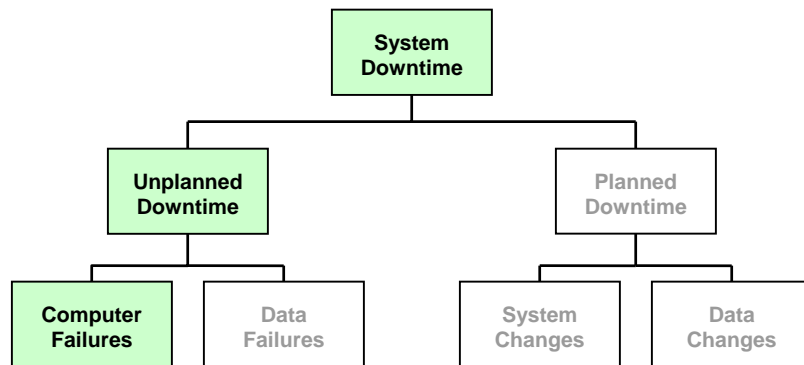


IT organizations that understand the different factors responsible for service interruption are better equipped to prevent outages. Through this understanding, robust high availability architectures can be implemented that are designed to protect against all causes of system downtime. In the following sections we will describe various Oracle Database technologies that can provide comprehensive protection against each of the failures mentioned above.

COMPUTER FAILURE PROTECTION

A computer failure is encountered when the machine running the database server unexpectedly fails, most likely due to hardware breakdown. This is one of the most common types of failures. Oracle Real Application Clusters, which is the foundation of Oracle's Grid Computing architecture, can provide the most effective protection against such failures.

Figure 2: Hardware Failures



Oracle Real Application Clusters (RAC) is the premier Grid Computing technology to maximize the availability, performance, and scalability of enterprise applications.

Real Application Clusters

Oracle Real Application Clusters (RAC) is the premier database clustering technology that allows two or more computers (also referred to as “nodes”) in a cluster to concurrently access a single shared database. This effectively creates a single database system that spans multiple hardware systems yet appears to the application as a single unified database. This extends tremendous availability and scalability benefits to all of your applications, such as:

- Fault tolerance within the cluster, especially computer failures.
- Flexibility and cost effectiveness in capacity planning, so that a system can scale to any desired capacity on demand and as business needs change.

Real Application Clusters enables enterprise Grids. Enterprise Grids are built out of large configurations of standardized, commodity-priced components: processors, servers, network, and storage. RAC is the only technology that can harness these components into useful processing systems for the enterprise. Real Application Clusters and the Grid dramatically reduce operational costs and provide new levels of flexibility so that systems become more adaptive, proactive, and agile. Dynamic provisioning of nodes, storage, CPUs, and memory allow service levels to be easily and efficiently maintained while lowering cost still further through improved utilization. In addition, Real Application Clusters is completely transparent to the application accessing the RAC database, thereby allowing existing applications to be deployed on RAC without requiring any modifications.

A key advantage of the RAC architecture is the inherent fault tolerance provided by multiple nodes. Since the physical nodes run independently, the failure of one or more nodes will not affect other nodes in the cluster. Failover can happen to any node on the Grid. In the extreme case, a Real Application Clusters system will still provide database service even when all but one node is down. This architecture allows a group of nodes to be transparently put online or taken off-line, for maintenance, while the rest of the cluster continues to provide database service. RAC provides built in integration with Oracle Fusion Middleware for failing over connection pools. With this capability, an application is immediately notified of any failure rather than having to wait tens of minutes for a TCP timeout to occur. The application can immediately take the appropriate recovery action. And Grid load balancing will redistribute load over time.

There is no better way to protect your application against server failures. Applications running on Real Application Clusters Database will continue to run even when all but one machines in the cluster is down.

RAC provides flexible scalability through dynamic hardware resource allocation. The capability to add hardware resources on-demand dramatically reduces IT costs allowing the IT infrastructure to grow based on business demand.

Real Application Clusters also gives users the flexibility to add nodes to the cluster as the demands for capacity increases, scaling the system incrementally to save costs and eliminating the need to replace smaller single node systems with larger ones. It makes the capacity upgrade process much easier and faster since one or more nodes can be incrementally added to the cluster, compared to replacing existing systems with new and larger nodes to upgrade systems. The Cache Fusion technology implemented in Real Application Clusters and the support for InfiniBand networking enables capacity to be scaled near linearly without making any changes to your application.

Oracle Database 11g further optimizes the performance, scalability and failover mechanisms of Real Application Clusters to further enhance its scalability and high availability benefits.

For more information on Real Application Clusters, please visit <http://www.oracle.com/technology/products/database/clustering/index.html>.

Bounding Database Crash Recovery Time

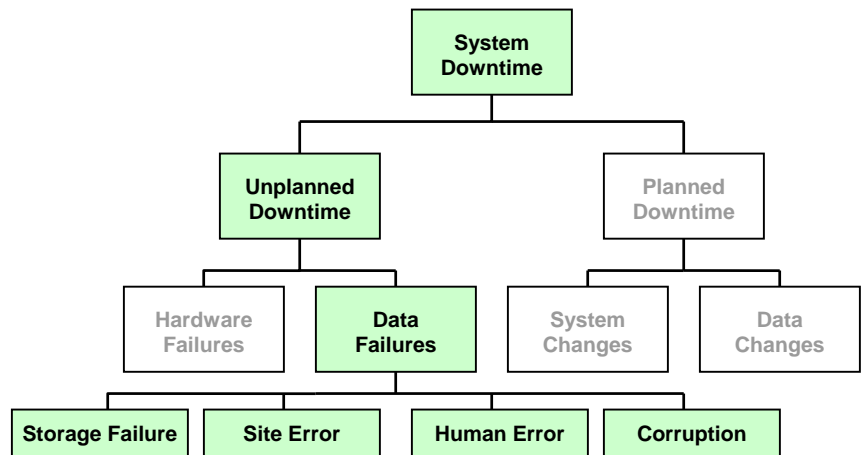
One of the most common causes of unplanned downtime is a system fault or crash. System faults are the result of hardware failures, power failures, and operating system or server crashes. The amount of disruption these failures cause will depend upon the number of affected users, and how quickly service is restored. High availability systems are designed to quickly and automatically recover from failures, should they occur. Users of critical systems look to the IT organization for a commitment that recovery from a failure will be fast and will take a predictable amount of time. Periods of downtime longer than this commitment can have direct effects on operations, and lead to lost revenue and productivity.

The Oracle Database provides very fast recovery from system faults and crashes. However, equally important to being fast is being predictable. The Fast-Start Fault Recovery technology included in the Oracle Database automatically bounds database crash recovery time and is unique to the Oracle Database. The database will self-tune checkpoint processing to safeguard the desired recovery time objective. This makes recovery time fast and predictable, and improves the ability to meet service level objectives. Oracle's Fast-Start Fault Recovery can reduce recovery time on a heavily loaded database from tens of minutes to less than 10 seconds.

DATA FAILURE PROTECTION

Data failure is the loss, damage, or corruption of business critical data. The causes of data failure are multifaceted and in many cases data failure can be illusive and difficult to identify. Generally, one or a combination of the following causes data failure: storage subsystem failure, site failure, human error, and/or corruption.

Figure 3: Data Failures



Storage Failure Protection

Oracle Database 10g introduced Automatic Storage Management (ASM), a breakthrough storage technology that integrates file system and volume manager capabilities specifically designed for Oracle database files. Through its low cost, ease of administration, and high performance characteristics ASM quickly became the storage technology of choice for IT administrators managing both stand-alone and RAC databases.

With performance and high availability as a primary objective, ASM builds on the principle of stripe and mirror everything. Intelligent mirroring capabilities allow administrators to define 2 or 3 way mirrors for the ultimate protection of critical business data. When disk failures occur, system downtime is avoided by utilizing the data available on the mirrored disks. If the failed disk is permanently removed from ASM, the underlying data is striped or rebalanced across the remaining disks to continue delivering high performance.

ASM Block Repair

Oracle Database 11g introduces new functionality to increase the reliability and availability of ASM. The first of these features is the capability to recover corrupt blocks on a disk by leveraging the valid blocks available on the mirrored disk(s). When a read operation identifies that a corrupt block exists on disk, ASM automatically relocates the bad block to an uncorrupted portion of the disk. In addition, administrators can now utilize the ASMCMD utility to manually relocate specific blocks due to underlying corruption of the disk.

With Oracle Database 11g, databases using ASM have increases in availability with the ability to perform rolling upgrades of their ASM instances.

Rolling Upgrades of ASM

ASM in Oracle Database 11g enhances the availability of the entire cluster environment with the capability to perform *Rolling Upgrades* of the ASM Software. ASM Rolling Upgrades permit administrators to keep their applications online while they upgrade ASM on individual nodes by keeping the other nodes in the cluster available during the migration. The ASM instances can run at different software versions until all nodes in the cluster have been upgraded. Any functionality introduced in the newer version of the ASM Software would not be enabled until all nodes in the cluster are upgraded.

Site Failure Protection

Enterprises need to protect their critical data and applications against catastrophic events that can take an entire data center offline. Events such as natural disasters and power and communication outages are a few examples of scenarios that can have detrimental effects on the data center. The Oracle Database offers a variety of data protection solutions that can safeguard an enterprise from costly downtimes due to complete site failures. The most basic form of protection is the off-site storage of database backups. While integral to an overall HA strategy, the process of restoring backups in a site-wide disaster can take more time than the enterprise can afford and the backups may not contain the most up to date versions of data. A more expeditious and comprehensive solution is to manage one or more duplicate copies of the production database in physically separate data centers.

Data Guard

Oracle Data Guard should be the foundation of every IT infrastructure's disaster recovery implementation. Data Guard provides the technology for deploying and managing one or more standby copies of a production database either in the local data center or in a remote data center, which could be located anywhere in the world. A variety of configurable options are available in Data Guard that allow administrators to define the level of protection they require for their business. Data Guard also works transparently across Grid clusters as the servers can be added dynamically to the standby database in the event a failover is required. Data Guard supports two types of standby databases – Physical Standby databases that use Redo Apply technology and Logical Standby databases that use SQL Apply technology.

Data Guard Redo Apply (Physical Standby)

A Physical Standby database is maintained and synchronized with the production database via the Redo Apply technology. The redo data of the production database is shipped to the Physical Standby, which using media recovery applies the changes from redo data to the standby database. Using Redo Apply, the standby database remains physically identical to the production database. Physical standby databases are good for providing protection from disasters and data errors. In the event of an error or disaster, the physical standby can be opened, and be used to provide data

services to applications and end-users. Because the efficient media recovery mechanism is used to apply changes to the standby database, it is supported with every application, and can easily and efficiently keep up with even the largest transaction workloads.

One of the key distinguishing features of Oracle's High Availability strategy is our relentless focus on making the high availability infrastructure fully useable from a day-to-day perspective. This allows customers to make productive use of their disaster recovery investment for a wide range of operations, such as offloading reporting workload or backup activities to the standby database or using the standby database for testing activities.

Physical Standby databases can be opened in read-only mode – even while redo data is continuously applied.

Physical Standby databases have always had the ability to be opened read-only, providing a means to offload production workloads that only require read access to the database. Historically, the drawback to this approach was the requirement that media recovery be quiesced while the Physical Standby database was opened in read-only mode; thus causing the Physical Standby database to become out of synch with the production database. Groundbreaking advancements in Oracle Database 11g allow media recovery to continue while the Physical Standby database is opened in read-only mode. This exciting new capability, called Physical Standby with Real Time Query, removes the aforementioned drawbacks of opening standby for read-only activity – now the Physical Standby database remains in synch with the production database even as it services read-only applications.

A key benefit of having a standby database that is physically identical to the production database is the ability to utilize this standby database as the source for backup activities. Oracle Database 10g introduced Block Tracking technology that keeps a log of which blocks have changed since the last incremental backup was performed and dramatically reduces the time required for incremental backups. Prior to Oracle Database 11g, the fast incremental backups using the block tracking technology could only be performed on the primary database. This restriction has been lifted in Oracle Database 11g allowing customers to offload all of their backup activities to the standby database.

Oracle Database 11g also introduces a new functionality called “Snapshot Standby” that allows a physical standby to be opened for read-write activities temporarily for testing activities without losing disaster protection. Using this functionality, a physical standby database is temporarily converted into a “snapshot standby” database that can be opened read-write to process transactions that are independent of the primary database for test or other purposes. A snapshot standby database will continue to receive and archive updates from the primary database, however, redo data received from the primary will not be applied until the snapshot standby is converted back into a physical standby database and all updates that were made while it was a snapshot standby are discarded. This enables production data to remain in a protected state at all times.

Finally, Oracle Database 11g can apply changes on the standby database in parallel thereby dramatically improving performance.

Data Guard SQL Apply (Logical Standby)

A Logical Standby database is maintained and synchronized with the production database via the SQL Apply technology. Rather than using media recovery to apply changes from the production database, SQL Apply transforms the redo data into SQL transactions and applies them to a database that is open for read/write operations. The ability to have the database open allows the Logical Standby database to be used concurrently to offload certain workloads from the production database. Many organizations leverage the Logical Standby for Reporting and Decision Support Systems that can be optimized by adding additional indexes and/or Materialized Views to the standby.

The SQL Apply process maintains the data integrity between the production and Logical Standby database by comparing the before-change values of the primary's redo data and the before-change values on the standby to avoid logical corruptions. The Logical Standby database therefore, is most importantly a data protection feature that ensures high availability with extended capabilities enhancing the scalability of the IT infrastructure.

Enhancements in Oracle Database 11g broaden the capabilities of logical standby databases, dramatically improve the apply performance and make it easier to use. In Oracle Database 11g, SQL Apply continues to add support for additional data types, other Oracle features, and PL/SQL, including:

- XMLType data type (when stored as CLOB)
- Ability to execute DDL in parallel on a logical standby database
- Transparent Data Encryption (TDE)
- DBMS_FGA (Fine Grained Auditing)
- DBMS_RLS (Virtual Private Database)

Data Guard Broker

The primary and standby databases, as well as their various interactions, may be managed by using SQL*Plus™. For easier manageability, Data Guard also offers a distributed management framework called the Data Guard Broker, which automates and centralizes the creation, maintenance, and monitoring of a Data Guard configuration. Administrators may use either Oracle Enterprise Manager or the Broker's own specialized command-line interface (DGMGRL) to take advantage of the Broker's management capabilities. From the easy to use GUI in Oracle Enterprise Manager, a single mouse click can initiate failover processing from the primary to either type of standby database. The Broker and Enterprise Manager make it easy for the DBA to manage and operate the standby database. By

facilitating activities such as failover and switchover, the possibility of errors is greatly reduced.

Oracle Database 11g further enhances Data Guard Broker to provide improved support for network transport option, eliminate downtime while changing the protection configuration (from Maximum Availability and Maximum Performance) and add support for single instance databases configured for HA using Oracle Clusterware as a cold failover cluster.

Fast-Start Failover

Data Guard Fast-Start Failover enables the creation of a fault tolerant standby database environment by providing the ability to totally automate the failover of database processing from the production to standby database without any human intervention. In the event of a failure, Fast-Start Failover will automatically, quickly, and reliably failover to a designated, synchronized standby database, without requiring administrators to perform complex manual steps to invoke and implement the failover operation. This greatly reduces the length of an outage. After a Fast-Start Failover occurs, the old primary database, upon reconnection to the configuration, will be automatically reinstated as a new standby database by the Broker. This enables the Data Guard configuration to restore disaster protection in the configuration easily and quickly, improving the robustness of the Data Guard configuration. Thanks to this feature, Data Guard not only helps maintain transparent business continuity, but also reduces the management costs for the DR configuration.

The new enhancements to Fast-Start Failover mechanism in Oracle Database 11g further reduce the failover time and provide administrators more control over the failover scenarios and behavior. For instance, Administrators can now define specific events, such as database errors (ORA-xxxx), which will trigger a Fast-Start Failover. Similarly, administrators can configure their Data Guard environment to shutdown the primary database when Fast-Start Failover is initiated in order to prevent accidental updates.

Human Error Protection

Almost any research done on the causes of downtime identifies human error as the single largest cause of downtime. Human errors like: the inadvertent deletion of important data; or when an incorrect WHERE clause in an UPDATE statement updates many more rows than were intended; need to be prevented wherever possible, and undone when the precautions against them fail. The Oracle Database provides easy to use yet powerful tools that help administrators quickly diagnose and recover from these errors, should they occur. It also includes features that allow end-users to recover from problems without administrator involvement, reducing the support burden on the DBA, and speeding recovery of the lost and damaged data.

Oracle automates the failover process through the use of the Fast-Start Failover feature.

Fast-Start Failover reduces the dependency of administrator availability to activate the standby in the event of a disaster.

Guarding Against Human Errors

The best way to prevent errors is to restrict a user's access to data and services they truly need to conduct their business. The Oracle Database provides a wide range of security tools to control user access to application data by authenticating users and then allowing administrators to grant users only those privileges required to perform their duties. In addition the security model of Oracle Database provides the ability to restrict data access at a row level, using the Virtual Private Database (VPD) feature, further isolating users from data they do not need access to.

Oracle Flashback Technology

When authorized people make mistakes, you need the tools to correct these errors. Oracle Database 11g provides a family of human error correction technology called Flashback. Flashback revolutionizes data recovery. In the past, it might take minutes to damage a database but hours to recover it. With Flashback, the time to correct errors equals the time it took to make the error. It is also extremely easy to use and a single short command can be used to recover the entire database instead of following some complex procedure. Flashback provides a SQL interface to quickly analyze and repair human errors. Flashback provides fine-grained surgical analysis and repair for localized damage -- like when the wrong customer order is deleted. Flashback also allows for correction of more widespread damage yet does it quickly to avoid long downtime -- like when all of this month's customer orders have been deleted. Flashback is unique to the Oracle Database and supports recovery at all levels including the row, transaction, table, tablespace, and database wide.

Flashback Query

Using Oracle Flashback Query, administrators are able to query any data at some point-in-time in the past. This powerful feature can be used to view and reconstruct logically corrupted data that may have been deleted or changed inadvertently.

```
SELECT *
FROM emp
AS OF TIMESTAMP
TO_TIMESTAMP('01-APR-07' 02:00:00 PM','DD-MON-YY HH:MI:SS PM')
WHERE ...
```

This simple query displays rows from the emp table as of the specified timestamp. This feature is a powerful tool that administrators can leverage to quickly identify and resolve logical data corruption. However, this functionality could easily be built into an application to provide application users with an easy and quick mechanism to rollback or undo changes to data without contacting their administrator.

Flashback Versions Query

Flashback Versions Query, similar to Flashback Query, is a feature that enables administrators to query any data in the past. The difference and the power behind Flashback Versions Query is its ability to retrieve different versions of a row across a specified time interval.

```
SELECT *
FROM emp
VERSIONS BETWEEN TIMESTAMP
  TO_TIMESTAMP('01-APR-07' 02:00:00 PM,'DD-MON-YY HH:MI:SS PM')
  AND
  TO_TIMESTAMP('01-APR-07' 03:00:00 PM,'DD-MON-YY HH:MI:SS PM')
WHERE ...
```

This query displays each version of the row between the specified timestamps. The administrator will have visibility into the values as they were modified by different transactions throughout this period. This mechanism gives the administrator the ability to pinpoint exactly when and how data has changed, providing tremendous value in both data repair and application debugging.

Flashback Transaction

Often times, a logical corruption can occur throughout a transaction that may change data in multiple rows or tables. Flashback Transaction Query allows an administrator to see all the changes made by a specific transaction.

```
SELECT *
FROM FLASHBACK_TRANSACTION_QUERY
WHERE XID = '000200030000002D'
```

Not only will this query show the changes made by this transaction, but it will also produce the SQL statements necessary to flashback or undo the transaction. A precision tool such as this empowers the administrator to delicately and efficiently diagnose and resolve logical corruptions in the database.

Flashback Transaction, new in Oracle Database 11g, is a seamless and powerful set of PL/SQL interfaces that simplify transaction-level data recovery. Building on the power of Flashback Transaction Query, this new feature enables a more robust and failsafe approach to repairing logical data corruptions. Many times, data failures can take time to be identified. When this is the case, it is possible that additional transactions have been executed based on logically corrupted data. Flashback Transaction identifies and resolves not only the initial transaction but all dependent transactions as well.

Flashback Data Archive, new to Oracle Database 11g, is a mechanism for storing historical versions of data for extended periods of time.

Flashback Data Archive

The Flashback query statements discussed above depend on the availability of the historical data in the UNDO tablespace. The amount of time that historical data remains in the UNDO tablespace is dependent on the size of the tablespace, the rate of data changes, and configurable database settings. Typically, administrators configure their databases to keep UNDO data no longer than days or weeks – certainly not years or decades. To overcome this limitation, Oracle Database 11g introduces pioneering new capabilities available through Flashback Data Archive. Flashback Data Archive maintains historical versions of data as regular data within the database that can be maintained for as long as required by the business. Flashback Data Archive revolutionizes data retention strategies to assist enterprises in the ever-changing regulatory landscape, such as Sarbanes-Oxley and HIPPA. To ensure the integrity of the retained data – Flashback Data Archive allows read-only access to the historical versions of data.

Automatically managed by Oracle, each time data is changed a read-only copy of the original version of data becomes available in the Flashback Data Archive.

The Flashback Data Archive is a robust tool-set that provides enterprises with amazing flexibility in managing their critical business data. Clearly, the advantages of Flashback Data Archive far surpass just the implicit benefits of repairing data failures. Using this technology, application developers and administrators can enable users to track and view information evolution. Given the immutable nature of the Flashback Data Archive, enterprises gain a strategic and financial advantage in terms of data preservation for purposes such as auditing. Application developers can take advantage of the Flashback Data Archive by introducing rich features into their applications allowing users to view past versions of data – such as banking statements. Finally, application developers and administrators are no longer burdened with creating and maintaining custom logic to track changes to critical business data.

Flashback Database

To restore an entire database to a previous point-in-time, the traditional method is to restore the database from a RMAN backup and recover to the point-in-time prior to the error. With the size of databases growing, it can take hours or even days to restore an entire database.

Flashback Database is a new strategy for restoring an entire database to a specific point-in-time. Flashback Database uses flashback logs to essentially rewind the database to the desired time. Flashback Database, using the flashback logs, is extremely fast as it only restores blocks that have changed. Easy to use and efficient, Flashback Database can literally restore a database in a matter of minutes in comparison to several hours.

```
FLASHBACK DATABASE TO TIMESTAMP  
TO_TIMESTAMP('01-APR-07 02:00:00 PM','DD-MON-YY HH:MI:SS PM')
```

As you can see, no complicated recovery procedures are required and there is no need to restore backups from tape. Flashback Database drastically reduces the amount of downtime required for scenarios requiring a database restore.

Flashback Table

Often times logical corruption is quarantined to one or a set of tables, thus not requiring a restore of the entire database. Flashback Table is the feature that allows the administrator to recover a table, or a set of tables, to a specific point-in-time quickly and easily.

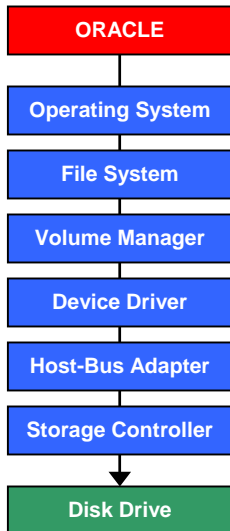
```
FLASHBACK TABLE orders, order_items TO_TIMESTAMP('01-APR-07 02:00:00 PM','DD-MON-YY HH:MI:SS PM')
```

This query will rewind the orders and order_item tables, undoing any updates made to these tables between the current time and the specified timestamp. In the event that a table is accidentally dropped, administrators can use the Flashback Table feature to restore the dropped table, and all of its indexes, constraints, and triggers, from the *Recycle Bin*. Dropped objects remain in the *Recycle Bin* until the administrator explicitly purges them or if the object's tablespace becomes pressured for free space.

Flashback Restore Points

In the above descriptions and examples of Flashback Database and Flashback Table, we have used time as the criteria for our restore or flashback operations. In Oracle Database 10g Release 2, Flashback Restore Points were provided as a means to simplify and expedite data failure resolution. A restore point is a user-defined label that *bookmarks* a specific time that the administrator believes the database to be in a good state. Flashback Restore Points allow administrators to more easily and efficiently remedy their databases from inappropriate and damaging activities.

IO Path



Data Corruption Protection

Physical data corruption is created by faults in any one of the various components making up the IO stack. At a high-level, when Oracle issues a write operation the database IO operation is passed to the operating system's IO code. This initiates the process of passing the IO through the IO stack where it is passed through the various components, from the file system to the volume manager to the device driver to the Host-Bus Adapter to the storage controller and finally to the disk drive where the data is written. Hardware failures or bugs in any one of these components could result in invalid or corrupt data being written to disk. The resulting corruption could damage internal Oracle control information or application/user data – either of which could be catastrophic to the functioning or availability of the database.

Through Oracle's unique HARD program, leading storage vendors implement Oracle's data validation algorithms directly in the storage device.

Oracle Hardware Assisted Resilient Data (HARD)

Oracle's Hardware Assisted Resilient Data is a comprehensive program that facilitates preventative measures to reduce the occurrences of physical corruption due to failures in the IO stack. This unique program is a collaborative effort between Oracle and leading storage vendors. Specifically, participating storage vendors implement Oracle's data validation algorithms within their storage devices. Unique to the Oracle database, HARD detects corruptions introduced anywhere in the IO path between the database and the storage device; this end-to-end data validation prevents corrupted data from being written to persistent storage. HARD has been enhanced to provide more comprehensive validation algorithms and support for all file types. Data files, online logs, archive logs and backups are all supported through the HARD program. Automatic Storage Management (ASM) utilizes the HARD capabilities without requiring the use of raw devices.

Backup and Recovery

Despite the power of the numerous preventative and recovery technologies discussed thus far in this paper, every IT organization must deploy a comprehensive data backup procedure. Scenarios when multiple failures occur at the same time, while rare, do happen and the administrator must be able to recover the business critical data from backup. Oracle provides industry standard tools to efficiently and properly backup data, restore data from previous backups, and to recover data up to the time just before a failure occurred.

Recovery Manager (RMAN)

Large databases can be composed of hundreds of files spread over many mount points, making backup up activities extremely challenging. Neglecting or overlooking even one critical file in a backup can render the entire database backup useless. As is too often the case, incomplete backups go undetected until they are needed in an emergency scenario. Oracle Recovery Manager (RMAN) is the composite tool that manages the database backup, restore, and recovery processes. RMAN maintains configurable backup and recovery policies and keeps historical records of all database backup and recovery activities. Through its comprehensive feature set, RMAN ensures that all files required to successfully restore and recover a database are included in complete database backups. Furthermore, through the RMAN backup operations, all data blocks are analyzed to ensure that corrupt blocks are not propagated throughout the backup files.

Oracle's Block Tracking technology, which greatly increases the speed of incremental backups, is now available for managed standby databases.

Enhancements to RMAN have made backing up large databases an efficient and straightforward process. RMAN takes advantage of Block Tracking capabilities to increase the performance of incremental backups. Only backing up blocks that have changed since the last backup vastly reduces the time and overhead of the RMAN backup. In Oracle Database 11g, the Block Tracking capabilities are now enabled on managed standby databases. With the size of enterprise databases continuing to grow – it has become more advantageous to take advantage of Bigfile Tablespaces. A Bigfile Tablespace is made up of a single large file rather than

numerous smaller files, allowing Oracle Databases to scale up to 8 exabytes in size. To increase the performance of backup and recovery operations of Bigfile Tablespaces – RMAN in Oracle Database 11g can perform intra-file parallel backup and recovery operations.

Many enterprises create clones or copies of their production databases to be used for testing, quality assurance, and to generate a standby database. RMAN has long had the capability to clone a database using existing RMAN backups via the DUPLICATE DATABASE functionality. Prior to Oracle Database 11g, the necessary backup files needed to be accessible on the host of the cloned database. Oracle Database 11g network-based duplication will duplicate the source database to the clone database without requiring the source database to have existing backups. Rather, the network-based duplication will transparently clone the necessary files directly from the source to the clone.

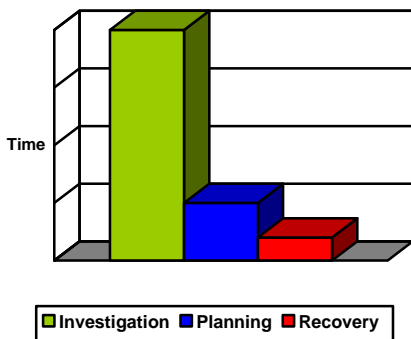
Oracle Database 11g supports a tight integration with Microsoft’s Virtual Shadow Copy Service (VSS). Briefly, Microsoft’s Virtual Shadow Copy Service is a technology framework that allows applications to continue to write to disk volumes while consistent point-in-time backups of those volumes are being performed. Oracle’s VSS Writer, a separate executable running as a service on Windows systems, will act as a coordinator between the Oracle database and other VSS components. For instance, the Oracle VSS Writer will put database files in hot backup mode to allow VSS components to take a recoverable copy of the data file in a VSS snapshot. The Oracle VSS Writer will leverage RMAN as the tool used to perform recovery on the files restored from a VSS snapshot. In addition, RMAN has been enhanced to utilize VSS snapshots as a source for incremental backups stored in the Flash Recovery Area.

Data Recovery Advisor

When the unthinkable situation arises and critical business data becomes jeopardized all recovery and repair options need to be evaluated to ensure a safe and fast recovery. These situations can be very stressful and often occur in the middle of the night. Research shows that administrators spend a majority of *Repair Time* performing investigation into what, why, and how data has become compromised. Administrators need to comb through volumes of information to identify the relevant errors, alerts, and trace files.

The Oracle Database 11g Data Recovery Advisor, built to minimize the time spent in the investigation and planning phases of recovery, reduces the uncertainty and confusion during an outage. Tightly integrated with other Oracle high availability features such as Data Guard and RMAN, the Data Recovery Advisor analyzes all recovery scenarios quickly and accurately. Through this integration, the advisor is able to identify which recovery options are feasible given the specific conditions. The possible recovery options are presented to the administrator, ranked based on recovery time and data loss. The Data Recovery Advisor can be configured to

Time to Repair



automatically implement the best recovery options, thus reducing any dependencies on the administrator.

Many disaster scenarios can be mitigated based on accurate analysis of errors and trace files that are presented prior to an outage. Therefore, the Data Recovery Advisor automatically and continuously analyzes the condition of the database through various health checks. As the advisor identifies symptoms that could be precursors to a database outage, the administrator can choose to obtain recovery advice and perform the necessary actions to fix the associated problem and avoid system downtime.

Oracle Secure Backup

Oracle Secure Backup, a centralized tape management system, backs up databases up to 25% faster than the leading competition.

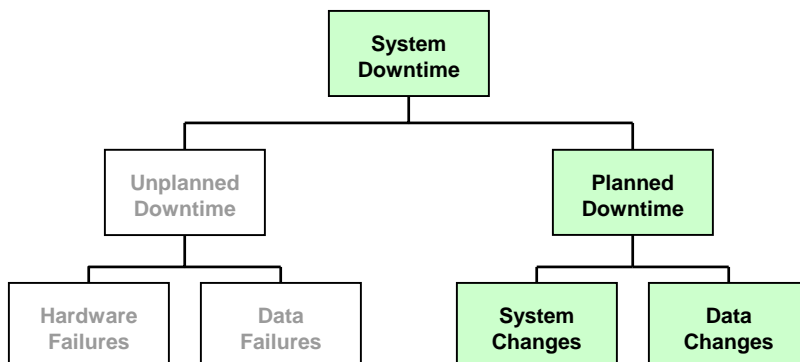
Oracle Secure Backup – a new product offering from Oracle – provides centralized tape backup management for entire Oracle environments including databases and file systems. Oracle Secure Backup offers customers a highly secure, cost effective and high performance tape backup solution. Thanks to its tight integration with Oracle Database, Oracle Secure Backup can back up an Oracle Database up to 25% faster than the leading competition. This is accomplished by leveraging direct calls into the database engine and through efficient algorithms that skip unused data blocks. This performance advantage will only continue to widen in the future as Oracle Secure Backup integrates even better with the database engine, thereby building special optimizations to improve backup performance even further.

Oracle Secure Backup is also integrated with Oracle Enterprise Manager – our web base GUI administrative tool – allowing administrators the unprecedented ease of use for setting up tape backups or restoring/recovering data from tape.

PLANNED DOWNTIME PROTECTION

Planned downtime is typically scheduled to provide administrators with a window to perform system and/or application maintenance. Throughout these maintenance windows, administrators take backups, repair or add hardware components, upgrade or patch software packages, and modify application components including data, code, and database structures. In today's networked global economy, enterprise applications and databases need to be accessible 24 hours a day. While advancements in networking and Internet technologies have had a profound impact on business productivity, these advancements have introduced new challenges and requirements for highly available architectures.

Figure 5: System Changes



Oracle has recognized administrator's need to continue traditional system and maintenance activities, while avoiding system and application downtime. Enhancements in Oracle Database 11g further promote this streamlined objective.

Online System Reconfiguration

Oracle supports dynamic online system reconfiguration for all components of your Oracle hardware stack. Oracle's Automatic Storage Management (ASM) has built-in capabilities that allow the online addition or removal of ASM disks. When disks are added or removed from an ASM Diskgroup – Oracle automatically rebalances the data across the new storage configuration while the storage, database, and application remain online. As discussed earlier in the paper, Real Application Clusters provide extraordinary online reconfiguration capabilities. Administrators can dynamically add and remove clustered nodes without any disruption to the database or the application. Oracle supports the dynamic addition or removal of CPUs on SMP servers that have this online capability. Finally, Oracle's dynamic shared memory tuning capabilities allow administrators to grow and shrink the shared memory and database cache online. With automatic memory tuning capabilities, administrators can let Oracle automate the sizing and distribution of shared memory per Oracle's analysis of memory usage characteristics. Oracle's extensive online reconfiguration capabilities support administrators' ability to not only minimize system downtime due to maintenance activities – but to also enable enterprises to scale their capacity on demand.

Online Patching and Upgrades

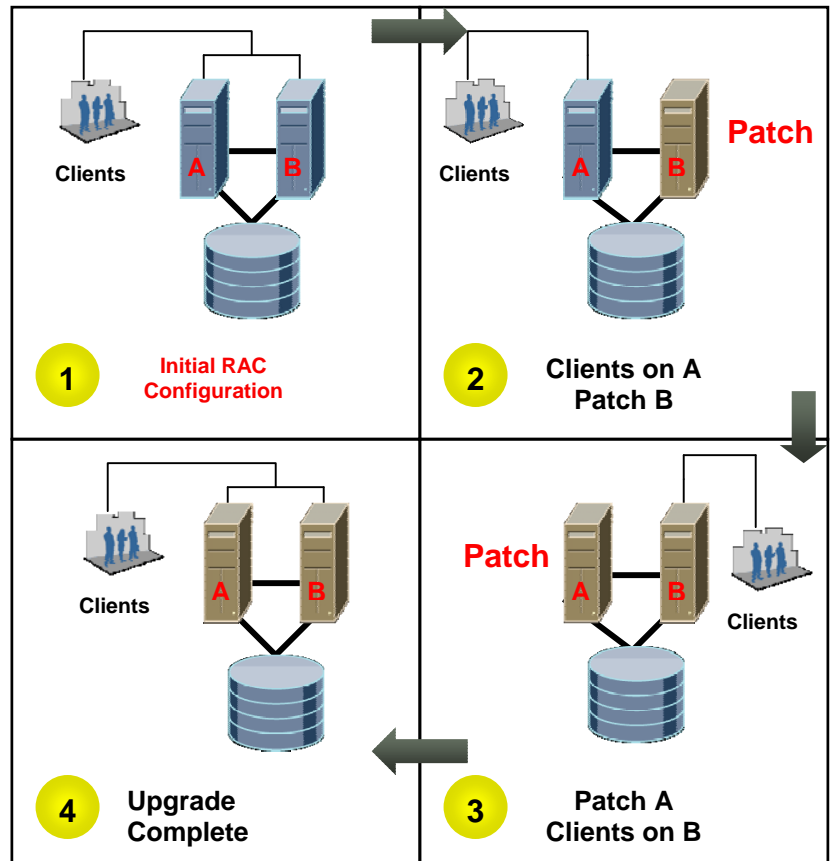
Enterprises with high availability demands can leverage Oracle technology to patch and upgrade their systems without end user interruption. With the strategic use of Real Application Clusters and Oracle Data Guard, administrators can more adeptly support the demands of the business.

Oracle's RAC and Data Guard features provide strategic capabilities to maintain application availability even during the application of patches, hardware maintenance, and software upgrades.

Rolling Patch Updates

Oracle supports the application of patches to the nodes of a Real Application Cluster (RAC) system in a rolling fashion permitting availability of the database throughout the patching process. The online patching process is illustrated in *Figure 6* below. The first box depicts a two node RAC cluster. To perform the rolling upgrade, one of the instances is quiesced while the other instance(s) in the cluster continue to service the end users. In the second box in our example, instance 'B' is quiesced and patched; meanwhile all client traffic is directed to instance 'A'. After the patch is successfully applied to the instance it can rejoin the cluster and be brought back online. Note that the instance(s) are now running at different maintenance levels and can continue to do so for an arbitrary amount of time. This allows the administrators to test and verify the newly patched instance before applying the patch to the rest of the instances in the cluster. Once the patch has been validated, the other instance(s) in the cluster can be quiesced and patched using the same rolling upgrade methodology. The third box in our example, illustrates instance 'A' being quiesced and patched and instance 'B' again accepting the client traffic. Finally, all instances in the cluster have been patched, are at the same maintenance patch level, and are again online balancing the client requests across the cluster. The rolling upgrade methodology can be used for emergency one-off database and diagnostic patches using OPATCH, operating system upgrades, and hardware upgrades.

Figure 6: Online Patch

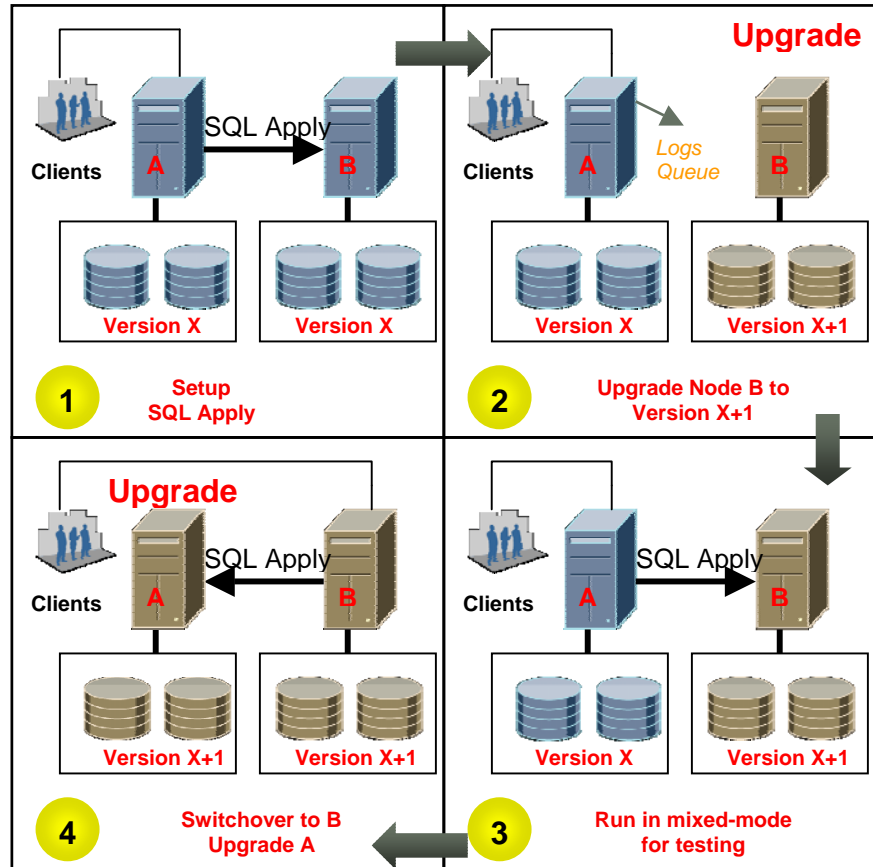


Online Software Upgrades

Utilizing Oracle's SQL Apply Data Guard technology, administrators can apply database patchsets, major release upgrades, and cluster upgrades with nearly no downtime to the end users. The process begins with instantiating a logical standby database and configuring Data Guard to keep the standby synchronized with the production database. Once the Data Guard configuration is complete, the administrator will pause the synchronization and all redo data will be queued. The standby database is upgraded, brought back online, and Data Guard is activated. All queued redo data will be propagated and applied on the standby to ensure no data loss occurs between the two databases. The standby and production databases can remain in *mixed-mode* until testing confirms the upgrade completed successfully. At this point, the switchover can occur resulting in a database role reversal – the standby database is now servicing the production workload and the production database is ready to be upgraded. While the production database is upgraded, the standby database (converted to primary during the switchover) is queuing the redo

data. Once the production database is upgraded and the redo data is applied, a second switchover takes place and the original production system is again taking production traffic. Figure 7 below illustrates the process for upgrading a database with near zero downtime.

Figure 7: Rolling Software Upgrade



Oracle Database 11g further enhances the appeal of the rolling upgrade process by introducing a functionality called “Transient Logical Standby”. This feature allows users to convert a physical standby to a logical standby database temporarily to effect a rolling database upgrade, and then revert to a physical standby once the upgrade is complete (using the `KEEP IDENTITY` clause). This benefits physical standby users who wish to execute a rolling database upgrade without investing in redundant storage otherwise needed to create a logical standby database.

Online Data and Schema Reorganization

Online data and schema reorganization improves the overall database availability and reduces planned downtime by allowing users full access to the database

throughout the reorganization process. Each release of Oracle has introduced enhanced online reorganization capabilities such as creating and rebuilding indexes, relocating and defragmenting tables, and adding, dropping, and renaming columns. Support of online reorganization functionality continues to be extended to additional object types including: advanced queuing (AQ) tables, materialized view logs, tables with Abstract Data Types (ADT), and Clustered Tables. Exciting new online reorganization functionality in Oracle 10g enabled administrators to reclaim unused space from segments – reducing the database footprint without end user interruption.

Additional improvements to online data and schema reorganization are being introduced in Oracle Database 11g. Traditionally, adding a column with a default value to a table with many rows could take a significant amount of time and essentially hold a lock on that table until the operation completed – inhibiting the availability of the application during this process. Advances in the method in which Oracle adds columns with default values has been significantly improved. Through these innovations, the overhead associated with the default value specification have been removed and therefore adding columns with default values have no impact on database availability nor performance.

Enhancements have been made to many data definition language (DDL) maintenance operations. Certain ddl operations are no longer forced to acquire NO WAIT locks. Administrators can define how long ddl operations are permitted to wait on locks before aborting the ddl operation. Many ddl operations have been enhanced to acquire sharing locks, rather than exclusive locks, throughout the duration of the maintenance operation. These advancements empower the administrator to maintain a highly available environment without impacting their ability to perform routine maintenance operations and schema upgrades.

Oracle Database 11g introduces a new attribute for indexes in order to increase availability throughout the schema maintenance and upgrade process. Indexes can now be created with the **Invisible** attribute causing the Cost-Based Optimizer (CBO) to ignore the presence of the index. Hints within SQL statements will make an invisible index ‘visible’ to the CBO, such that maintenance and upgrade SQL statements can leverage an index without causing application SQL to erroneously use an index. While the index is invisible to the CBO, invisible indexes are still maintained by DML operations. When an index is determined to be ready for production availability, a simple *Alter Index* statement will make the index visible to the CBO.

Application Upgrades

As business requirements evolve, so too do the applications and databases supporting the business. Historically, application upgrades necessitated planned downtime. Through the strategic use of the DBMS_REDEFINITION package (also available in Enterprise Manager) – administrators can seamlessly manage application upgrades while continuing to support an online production system.

Administrators using this API, enable end users to access the original table, including insert/update/delete operations, while the upgrade process modifies an interim copy of the table. The interim table is routinely synchronized with the original table and once the upgrade procedures are complete, the administrator performs the final synchronization and activates the upgraded table.

Partitioning

As databases grow, they can become more challenging to manage. Partitioning is a pivotal technology that allows administrators to break large tables and indexes into smaller, more manageable pieces. While most maintenance activities can be performed online, performing maintenance one partition at a time provides flexibility and performance benefits to most online operations. Furthermore, partitioning increases the fault tolerance of the Oracle Database. Administrators can strategically locate individual partitions on different disks; therefore a disk failure will only affect the partitions that reside on that disk.

MAXIMUM AVAILABILITY ARCHITECTURE – BEST PRACTICES

Operational best practices are essential to the success of an IT infrastructure. Oracle's Maximum Availability Architecture (MAA) is Oracle's best practices blueprint based on the integrated suite of Oracle's best-of-breed High Availability (HA) technologies. MAA integrates Oracle Database features for high availability including Real Application Clusters, Data Guard, Recovery Manager, and Enterprise Manager. MAA includes best practice recommendations for critical infrastructure components including servers, storage systems, network systems, and application servers. Beyond the technology, the MAA blueprint encompasses specific design and configuration recommendations that have been tested to ensure optimum system availability and reliability. Enterprises that leverage MAA in their IT infrastructure find they can quickly and efficiently deploy applications that meet their business requirements for high availability.

Oracle's Maximum Availability Architecture, through the right combination of technology and operational best practices, enables enterprises to deploy unbreakable IT solutions. The MAA best practices are continually being extended. For additional information regarding MAA please visit <http://otn.oracle.com/deploy/availability/htdocs/maa.htm>.

CONCLUSION

Enterprises understand the critical value in maintaining highly available technology infrastructures to protect critical data and information systems. At the core of many mission critical information systems is the Oracle database, responsible for the availability, security, and reliability of the technology infrastructure. Building on decades of innovation, Oracle Database 11g introduces revolutionary new

Oracle's Maximum Availability Architecture is the integration of best-of-breed technologies providing the most comprehensive and cost-effective suite of High Availability technologies.

availability and data protection technologies to provide customers with new and more effective ways of maximizing their data and application availability. Oracle's comprehensive set of technologies provides businesses unparalleled protection against any kind of outages – be it due to a planned maintenance activity or an unexpected failure. And the Grid capabilities provided make certain that the cost to deploy your database environment, and adapt to changing business needs, is significantly less than what you had to spend in the past to achieve equivalent results.



Oracle Database 11g High Availability
June 2007
Author: William Hodak
Contributing Author: Sushil Kumar

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com

Copyright © 2007, Oracle. All rights reserved.

This document is provided for information purposes only and the contents hereof are subject to change without notice.

This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission. Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.