MAA Best Practices
Replication Technologies for High Availability and Disaster Recovery – A Technical Overview

Joseph Meeks, HA\MAA Product Management, Oracle
Lawrence To, MAA Development, Oracle
Agenda

• The Role of Replication in HA/DR
• Challenges
• Oracle High Availability Replication Solutions
  • Active-Active Defined
  • Oracle Data Guard Physical & Logical Standby
  • Oracle Streams
  • Oracle RAC Extended Distance Clusters
• Oracle’s Strategic Direction
The Role of Replication in HA

Imagine... engine = database, plane = your business
The Role of Replication in HA

Imagine... engine = database, plane = your business
The Role of Replication in DR

Imagine your Data Center is in Toronto

14 August 2003
0129Z

~20 hrs before Blackout

[Map showing cities like Toronto, Montreal, Ottawa, Buffalo, and others, highlighting the geographical context.]
The Role of Replication in DR

Imagine your Data Center is in Toronto
Disaster Recovery Challenge

• Forrester Research, “Six Years After 9/11, Most Firms Are Not Ready For Another Disaster”, Sep 11, 2007

“… IT operations professionals are crossing their fingers and hoping a disaster won’t hit, while business executives have no idea how vulnerable they really are to significant losses . . .” (survey of 189 enterprises)
Even Companies Who Implement DR . . . using Traditional DR Technologies

• Are saddled with expensive, redundant systems
• Under-utilize their standby resources
• Have no immediate ROI until a disaster occurs
• Lack confidence that it will work when needed
Just Your Average “Disaster”

Financial Services Company

• Errors observed in the alert.log of the production database:
  
  • Errors in file /opt/app/oracle/admin/dg/bdump/dg1.trc:
  
  • ORA-01186 : file 93 failed verification tests
  
  • ORA-01122 : database file 93 failed verification check
  
  • ORA-01110 : data file 93: '/dbmnt/db01/oradata/dg/arch05.dg'
  
  • ORA-01251 : Unknown File Header Version read for file number 93

  • ORA-01251 - Corrupted file header. This could be caused due to missed read or write or hardware problem or process external to oracle overwriting the information in file header.

• Affected database:
  
  • Primary customer facing applications for trade transaction confirmation, new accounts, and customer account information

What Did They Do???
Challenges

Array Based Remote-mirroring

- Zero knowledge of the application or data
- Must mirror everything (all writes to all files)
- SYNC mode makes every DB write a synchronous write
- Mirrors problems just as effectively as it mirrors data
- The application (target Oracle DB) is off while mirroring is on
- Target volumes can not be used while mirroring is active
- Difficult to test – not sure if it will work when needed
- Vendor lock-in to costly storage subsystems
- Distance constraints

Limited Protection, Expensive, Low Return on Investment
More Challenges

3\textsuperscript{rd} party Transactional Replication Technologies

Attempt to address shortcomings of remote mirroring, however . . .

- Have limitations, e.g. data types, DDL
- Can not guarantee a zero data loss recovery point objective
- Deliver insufficient performance for high volume applications
- Are external to Oracle - complicates management & release upgrades
- Lack integration with other Oracle features
- Can be expensive to license
- Will impose costly lock-in to long-term 3\textsuperscript{rd} party support

\textit{Limited Protection, Expensive, More Complex to Support}
The Oracle Approach

Data-aware High Availability & Data Protection

• HA/DR replication solutions integrated with the database allow building intelligent capabilities
  • Data is always transactionally consistent
  • Data corruptions are detected and prevented
  • Assurance that data can be used after failover
  • Recovery at the level of business objects
  • Failover can be automatic, very fast and application aware
  • Agnostic of application, system vendor or storage
  • Protection from a wider variety of failures
  • Zero data loss protection
  • Reduction of planned downtime using database rolling upgrades
  • High performance
• All resources utilized productively
Oracle’s Integrated HA/DR Solution Set

Unplanned Downtime

- System Failures
- Data Failures

Planned Downtime

- System Changes
- Data Changes

Real Application Clusters

ASM
Flashback
RMAN & Oracle Secure Backup
Data Guard
Streams

Online Reconfiguration
Rolling Upgrades

Online Redefinition

Oracle MAA Best Practices
The Convergence of HA & DR

Oracle Maximum Availability Architecture (MAA)

Ensure data protection and availability by minimizing or eliminating downtime at all technology stack layers including hardware or software components while …

…protecting against any failure, including catastrophic failures caused to a data center or location because of earthquakes, fire, terrorist attacks, or floods.
So . . Back to that Real-life Disaster. . .

Financial Services Company

- Errors observed in the alert.log of the production database:
  - Errors in file /opt/app/oracle/admin/dg/bdump/dg1.trc:
  - ORA-01186: file 93 failed verification tests
  - ORA-01122: database file 93 failed verification check
  - ORA-01110: data file 93: '/dbmnt/db01/oradata/dg/arch05.dg'
  - ORA-01251: Unknown File Header Version read for file number 93

  - ORA-01251 - Corrupted file header. This could be caused due to missed read or write or hardware problem or process external to oracle overwriting the information in file header.

- They had already implemented Data Guard.
  - Corruptions were isolated to the primary
  - They quickly failed over to the standby & resumed operation with zero data loss
  - Users were blissfully ignorant of the three days spent by IT to resolve problem
Agenda

• The Problem
• Oracle High Availability Replication Solutions
  • Active-Active Defined
  • Oracle Data Guard Physical & Logical Standby
  • Oracle Streams
  • Oracle RAC Extended Distance Clusters
• Oracle’s Strategic Direction
# Active-Active Solutions for HA/DR

## Relative Comparison

<table>
<thead>
<tr>
<th>Technology</th>
<th>Standby Utilization</th>
<th>Complexity</th>
<th>Active Meter</th>
<th>DR Meter</th>
</tr>
</thead>
</table>
| Data Guard Physical Standby     | • Read-only queries and reporting  
• QA testing, backups                                                              | Lowest     | Very Good    | Best     |
| Data Guard Logical Standby      | • Add local tables, additional schemas, indexes and materialized views  
• Read-write access for reporting apps                                             | Moderate   | Better       | Very Good|
| Oracle Streams                  | • Advanced replication capabilities  
• Completely independent replicas  
• Run workloads anywhere                                                          | Highest    | Best         | Good     |
| Oracle RAC Extended Distance Cluster | • Load-balance workload across nodes  
• RAC, CRS and ASM redundancy eliminates third party components                    | Highest    | Best         | Limited  |
Apply Performance Comparison

- **Data Guard Physical Standby**
  - Apply performance scales up to system resources
    - I/O and CPU (MRP or recovery coordinator)
    - Observed apply rate for batch updates 50-100 MB/sec
    - Observed apply rate for OLTP type applications 10–60 MB/sec
- **Data Guard Logical Standby**
  - Array or batch DMLs translate to single row DML on standby
  - Additional tuning for loads, parallel DML, long running txns
  - Observed apply rate for LOB insert <= 30 MB/sec; for OLTP < 5 MB/sec;
    includes SQL Apply optimizations for large transactions
- **Oracle Streams**
  - Apply performance characteristics similar to Data Guard logical standby
  - Additional overhead: transformations, DML handlers, conflict resolution
  - Tuning is simplest with Physical Standby. More involved for others.
Oracle Data Guard
Physical & Logical Standby Databases
Oracle Data Guard

*Best Protection at Lowest Cost*

**Production Database**

**Automatic Failover**

**SYNC or ASYNC Redo Shipping**

**Physical or Logical Standby DB**

**Data Guard**
Data Guard Redo Transport

Network Traffic Dramatically Reduced Compared to Remote-mirroring

*www.oracle.com/technology/deploy/availability/htdocs/DataGuardRemoteMirroring.html*
Data Guard Apply Services

Multi-layer Oracle Validations Protect Standby from Corruptions

Updates

*www.oracle.com/technology/deploy/availability/htdocs/DataGuardRemoteMirroring.html*
Data Guard

Best Protection from Physical and Logical Corruptions

- Automatic validation detects redo corruptions
  - Checksum and internal meta-data checks for all reads and writes

- Ensures redo is “next change” appropriate to data block
  - Detects and prevents lost writes (11g) and data block corruptions from harming standby database (ultimate data protection)

- Validates data block during reads and after updates
  - Detects data block corruption through checksum during reads
  - Detects data block corruptions through db_block_checking after updates
Data Guard Failover
Minimizing Unplanned Downtime

• It is quicker to fail over than to diagnose and repair
• Data Guard can completely automate database failover
  • Eliminate human error and delay in responding to outage
• Flashback Database can quickly reinstate the old primary
• Use Data Guard switchover to frequently test DR readiness
• See MAA Best Practices for Switchover/Failover
• See MAA Best Practices for Fast-Start Failover

See: http://www.oracle.com/technology/deploy/availability/htdocs/maa.htm
Amazon.com Experience

Data Guard Fast-Start Failover

End-to-End Failover Time*

With Data Guard
Database failover: 20 secs
Apps redirected: 2 mins
Standby site distance: 15 miles

Before Data Guard

Data Guard Automatic Failover

Minutes

Resolve
Respond
Identify
Application Failover

Primary Site and Application Tier is Still Viable

Primary Site

- Database Services
- Primary Database

Standby Site

- Database Tier - Oracle Real Application Clusters
- Application Tier - Oracle Application Server Clusters

1. Data Guard Manual or Automatic Failover
2. Startup trigger is used to relocate primary database services
3. FAN breaks clients out of TCP timeout, applications quickly reconnect to new primary

- Standby becomes primary database

- MAA Best Practices for Client Failover in a Data Guard Configuration
- Demonstrations – Automatic Database, Application & Complete Site Failover
Data Guard Switchover

Minimizing Planned Downtime

- Data Center moves
  - Create standby at new data center then switch over
- Migrating to ASM and or RAC
  - Create ASM standby for non-ASM primary, then switch over, same for RAC
- Technology Refresh
  - Create standby database on new servers/storage then switchover
- Use physical standby for certain types of migrations *
  - Windows>Linux, 32bit>64bit, HP-UX/PA[-RISC] ->HP-UX/IPF
- Use physical standby for rolling database upgrades
  - By utilizing Transient Logical Standby in Data Guard 11g
- Use logical to implement database changes in rolling fashion
  - e.g. ASSM, Initiants, Blocksize – implement on standby & switchover
- Use logical standby for rolling database upgrades
  - MAA Best Practices for Rolling Database Upgrades

* see Metalink Note 413484.1
Data Guard
Physical Standby
Data Guard Physical Standby

Choose **Physical Standby** if:

1. Management simplicity is preferred
2. Highest performance is required
3. An exact replica of the primary is preferred for DR
4. Standby database also serves QA role - Snapshot Standby
5. A current, read-only standby can offload primary database*
6. The standby is used to offload backups from the primary
7. Protection from lost writes and corruptions is desired
8. Data type restrictions exclude using Logical or Streams

* Requires additional license for Oracle Active Data Guard
Active Data Guard

*Improve Primary Performance!*

- Offload read-only queries to an up-to-date physical standby
- Perform fast incremental backups on a physical standby

A “Read-Write” and a “Read Only” service are both running on the primary database.

An application having 50 users is connected to the Read-Write service.

A second application having 100 users is connected to the Read-Only service.
The “Read Only” service is stopped on the primary, and moved to the standby database (Paris 1).

Note the increase in TPS for each service.
Read-Write TPS increases 200%
Read-Only TPS increases 60%
Another instance of the The "Read Only" service is started to a second node on the Standby RAC cluster (Paris 2).

Read-Only TPS increases another 100%
Snapshot Standby

Combine Testing with DR

Physical Standby
Apply Logs

• Convert Physical to Snapshot Standby - open read-write to test applications
  • ALTER DATABASE CONVERT TO SNAPSHOT STANDBY;
• Discard testing writes and re-sync with primary by applying archived logs
  • ALTER DATABASE CONVERT TO PHYSICAL STANDBY;
• Preserves zero data loss
• No idle resources
• Similar to storage snapshots, but:
  • Provides DR at the same time
  • Uses single copy of storage
• Ideal QA system when combined with Real Application Testing

Snapshot Standby
Perform Testing

Physical Standby
Apply Logs

Open Database
Back out Changes

Continuous Redo Shipping
Rolling Database Upgrades

For physical standby databases

- Transient Logical Standby - Oracle11g
  - Temporarily convert physical standby to logical

  SQL> ALTER DATABASE RECOVER TO LOGICAL STANDBY KEEP IDENTITY

- Execute the upgrade
  - When upgrade is complete – revert to physical standby

- No need for separate logical standby
  - Physical standby users utilize existing servers and storage
Physical Standby

*Deployment Considerations*

- Physical standby deployment is SIMPLE
- Redo Apply is VERY FAST
- Flexible configuration options
- For aggressive Recovery Time Objectives
  - Follow MAA recovery and redo transport best practices
  - Follow MAA best practices to optimize switchover and failover
  - Consider automating database and client failover
- Offload primary resources and provide read scalability
- Use standby database for QA and other test activity

Mobiltel (M-Tel)

Oracle Database 11g Active Data Guard

Business systems

SOA

Resource Manager

Active Data Guard
DR with Real-Time Query

no idle resources

Continuous Redo
Shipment and Apply

Primary RAC Database

Standby Database Open Read-Only
Mobiltel (M-Tel)

Oracle Database 11g Snapshot Standby

Business systems

Capture Workload

RAC Primary Database

Snapshot Standby - Open Read-Write

SOA

Resource Manager

Replay Driver

Process
Data Guard
Logical Standby
Choose **Logical Standby** if:

- Flexibility of a logical replica (open read write) is preferred
  - Add local tables, additional schemas, indexes, materialized views, or other structures to optimize the standby database for other uses
  - You can offload applications that read and not update “primary” data

- Rolling database upgrade is required*

- Additional level of data protection due to logical replication

- Note – see prerequisites & best practices for logical standby
  - Chapter 4.1:Prerequisites - Data Guard Concepts & Admin
  - MAA Best Practices for SQL Apply

*Note: physical standby can be temporarily converted to logical standby for rolling database upgrade
**SQL Apply Process Architecture**

Redo Data from Primary Database

- **Reader**
- **Preparer**
- **Builder**

Logical Change Records not grouped into transactions

**Log Mining**

- LCR: Shared Pool

**Apply Processing**

- **Applier**
- **Coordinator**
- **Analyzer**

- **Transaction groups**

- **Transactions to be applied**
- **Transactions sorted in dependency order**
SQL Apply

*Extended Data Type Support*

- For tables containing the following unsupported types:
  - Object columns with simple or nested objects
  - Varrays
  - Partial Spatial types (SDO_GEOMETRY)
- When: planned for 10.2.0.4 and 11.1.0.7
- How: Customers can extend SQL Apply using table triggers and log tables to synchronize tables having these types
- Additional types planned for support in future releases
  - XML (object relational), object tables
SQL Apply – Rolling Database Upgrades

1. Initial SQL Apply Config
   - Clients
   - Redo
   - Version X
   - Version X

2. Upgrade node B to X+1
   - Logs
   - Queue
   - Upgrade
   - A
   - X
   - B
   - X+1

3. Run in mixed mode to test
   - Upgrade
   - A
   - B
   - X
   - X+1

4. Switchover to B, upgrade A
   - Upgrade
   - A
   - B
   - X+1
   - X+1

Patch Set Upgrades
Major Release Upgrades
Cluster Software & Hardware Upgrades
Logical Standby

Deployment Considerations

• Large performance improvements in every release
• Improved manageability and usability
• More data type restrictions being lifted in each release
• Key customer deployments and references
• Flexible configuration options
• For aggressive Recovery Time Objectives
  • SQL Apply and Redo Transport best practices
  • Optimize switchover and failover
  • Consider automating database and client failover (JDBC clients only)
• Offload primary database and provide reporting business value

Thomson Legal
SQL Apply: HA/DR, Rolling Upgrade & Tech Refresh

Primary Site

1. Create logical standby, C
2. Upgrade logical standby: 10.1.0.3 to 10.1.0.4, implement db changes
3. Create physical standby, D
4. Switchover to new logical standby, C becomes new primary – zero downtime for read-only access, very short downtime for inserts, updates, and deletes
5. Convert physical to logical
6. Decommission old platforms A & B

Secondary Site

Note: Thomson presented this process at OpenWorld 2006
Oracle Streams
Oracle Streams Replication

- All sites active and updateable
- Automatic conflict detection & optional resolution
- Supports data transformations
- Flexible configurations – n-way, hub & spoke, …
- Database platform / release / schema structure can differ
- Provides HA for custom apps where update conflicts can be avoided or managed
Oracle Streams

Choose Oracle Streams if:

• Concurrent updates on the same data in multiple sites
• For more than simple, one-way replication architectures
  • N-way multimaster (bi-directional replication), hub & spoke, many to one replication
  • Support data transformations, subsetting, custom apply functions
  • Support heterogeneous platforms and different charactersets
• Additional level of data protection due to logical replication
• To reduce planned downtime during application upgrades, when the logical structure of the database is changed
Streams Process Architecture

Capture at Source, Downstream, or Target Database

Apply at Target Database

Redo Data
Primary Database

Reader → Preparer → Builder → Capture

Logical Change
Records not grouped into transactions

LCR LCR:
Streams Pool

Propagation

Target Database
Datafiles

Applier → Coordinator → Reader

Conflict Detection
Error Handling
Custom Code

Committed transactions grouped and sorted in dependency order

Transactions to be applied
Streams Capture and Apply

Replication Configuration

Update EMP set job='coding' where empid=510;

Local (Upstream) Capture
If subset of data, network bandwidth requirements reduced significantly
Downstream Capture

Update EMP set job='coding' where empid=510;

Redo Transport Method ASYNC or SYNC

Support for Zero Data Loss
Streams

Deployment Considerations

• Improved manageability, usability, and performance in every release
• Extremely flexible architecture that allows for:
  • N-way multimaster and more sophisticated distributed environments
  • Data transformations, subsetting, custom apply functions
  • More heterogeneous platforms and different character sets
  • Instantaneous availability when site/database goes down
• Same extended data type support as SQL Apply (10.2.0.4 & 11.1.0.7)
  • Object columns (simple or nested objects), Varrays, partial spatial types
  • Utilizes a DML handler or trigger on target database
• Additional operational investment to
  • Avoid and handle conflicts for update anywhere architectures
  • Handle distributed recovery and data loss scenarios
  • Implement disaster recovery or rolling application upgrade solutions

Best Practices: Streams documentation and future papers in MAA OTN site
Volkswagen AG

Oracle Streams Hub and Spoke Configuration

- 1 TB central engineering repository (hub) maintained at German HQ
- Replicas (spokes) at factories worldwide for fast, reliable, local access
- Replicas synchronized with Streams
- Data Guard physical standby for protecting central repository

- $100,000 - $200,000/hr per site savings in downtime costs
- Fully bi-directional, automatic conflict detection and resolution
- Minimum subset of data replicated across WAN (about 1/3 out of 200 tables)

Oracle RAC
Extended Distance Clusters
Oracle RAC on Extended Clusters

- Full utilization of resources
- Fastest possible recovery from site failure

Work Continues on Remaining Site

Site A

One Physical Database

Site B
Oracle RAC

Choose Extended Distance Clusters if:

• Highest level of availability for server and site failures
• When Data Centers are in close enough proximity
  • Distance restrictions limit protection from regional failures

Caution: Extended RAC provides the highest possible level of HA, but by itself is NOT an HA/DR solution
One Database Cluster
Two Data Centers

Data Center A

Public Network

(Images of servers and network components)

Data Center B

Public Network

(Images of servers and network components)

RAC Network

Storage Network

RAC VLAN
Storage VLAN

10 Gbe

Metro Cluster

Database

ASM Internal Mirroring (sync between failure groups)

Preferred Mirror Read (Oracle 11g)
Extended RAC

Design Considerations

• Connectivity
  • Redundant connections and sufficient bandwidth for public traffic, interconnect, and I/O
  • High interconnect and network latency can throttle database performance and response time
  • Distance > 10km may require Dark Fiber which can incur higher costs
• Disk Mirroring to maintain synchronous copy at each site
  • ASM recommended, particularly for Oracle 11g
  • Do not use array-based mirroring because storage updates occur only on one “master” site (SPOF), longer recovery time, possible performance implications and more complex technology stack
• Requires 3rd site to host quorum device
• Vulnerable to data corruptions, lost writes or database wide failures
  • Still a single physical database
Extended RAC

Deployment Considerations

• Improved manageability and usability in every release
• Full utilization of all system and database resources
• Highest level of availability for server and site failures
• Data is always current (single physical database)
• However
  • NO DR, NO full data protection and NO database rolling upgrade
  • Distance limitations since it’s a single physical database
• Complement with a remote Data Guard Standby Database
  • Maintains an independent physical replica of production
  • Protection from regional disasters
  • Protection against corruptions and other potential failures
  • Provides option for rolling database upgrades and patch-sets
 Agenda

• The Problem
• Oracle High Availability Replication Solutions
  • Active-Active Defined
  • Oracle Data Guard Physical & Logical Standby
  • Oracle Streams
  • Oracle RAC Extended Distance Clusters
• Oracle’s Strategic Direction
Oracle Strategic Direction

HA/DR Replication Technologies

• Data Guard Physical Standby
  • More options for utilization of a physical standby database
  • Faster, even more transparent database and application failover
  • Offload and scale reads to standby

• Logical Standby & Oracle Streams
  • Improved performance, fewer restrictions, better manageability
  • Logical will always focus on simpler, one-way replication features
  • Logical offload reporting. Streams provide bi-directional updates.
  • Streams will always focus on sophisticated replication capabilities

• Oracle RAC – Extended Distance Clusters
  • Will benefit from general RAC (GRID) and ASM enhancements.
  • Data is always current
  • Best DR architecture includes remote Data Guard standby database
Oracle Strategic Direction

**HA/DR Replication Technologies**

- Related Enhancements
  - Further options for higher utilization and application integration
  - Further development of integrated, full stack replication
  - Further automation of complete site failover
  - Further reduction in downtime required for planned maintenance
  - Further improvements on manageability and ease of use
Resources

- **Oracle.com**
  http://www.oracle.com/ha

- **Maximum Availability Architecture Best Practices**
  http://www.oracle.com/technology/deploy/availability/htdocs/maa.htm

- **Oracle HA Customer Success Stories on OTN**

- **Article “Application-aware data protection: Databases”, InfoStor, Nov 2007**
## Which Solution To Use

<table>
<thead>
<tr>
<th>IF:</th>
<th>THEN</th>
</tr>
</thead>
</table>
| • Very high primary database throughput  
• Simplicity of a physical replica  
• Maximum protection from physical corruptions  
• Read-only access to synchronized standby can offload production | Data Guard Physical Standby |
| • Simpler, one-way logical replication  
• Standby requires local tables, additional schemas, indexes and mv’s  
• Offload apps that need read-write access, but don’t modify primary data | Data Guard Logical Standby |
| • Fine grained, N-way multimaster, hub&spoke or many to one replication  
• Load balance update workload across sites  
• Accommodations in app design and management are acceptable in return for access to advanced replication features | Oracle Streams |
| • Highest level of availability for server or computer room failure  
• HA benefit and workload balancing outweighs performance concerns  
• Willing to make additional provisions for remote data protection | Oracle RAC Extended Distance Cluster |
The preceding 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.