Zero Data Loss Recovery Appliance
Best Practices at Work in Your Data Center

Andrew Babb
Consulting Member of Technical Staff
MAA Development, System Technology Group

Dongwook Kim
Engineering System Team, Oracle Korea

JeongRyun Park
IT Planning Team Leader / Information Technology Office
SK Hynix

Modified Version of Oracle OpenWorld Presentation November 23, 2015
Zero Data Loss Recovery Appliance Overview

Protected Databases

**Delta Push**
- DBs access and send only changes
  - Minimal impact on production
  - Real-time redo transport instantly protects ongoing transactions

Protected all DBs in Data Center
- Petabytes of data
- Oracle 10.2-12c, any platform
- No expensive DB backup agents

Recovery Appliance

**Delta Store**
- Stores validated, compressed DB changes on disk
- Fast restores to any point-in-time using deltas
- Built on Exadata scaling and resilience
- Enterprise Manager end-to-end control

Offloads Tape Backup

Replicates to Remote Recovery Appliance
Program Agenda

1. Architecture
2. Backup Best Practices
3. Restore and Recovery
4. Recovery Appliance and Data Guard
5. Validation, Security and Troubleshooting
6. Best Practices at work
7. SK Hynix Recovery Appliance Use Case
Program Agenda

1. Architecture
2. Backup Best Practices
3. Restore and Recovery
4. Recovery Appliance and Data Guard
5. Validation, Security and Troubleshooting
6. Best Practices at work
7. SK Hynix Recovery Appliance Use Case
Architecture - Networks

• There are **three** networks on a RA
  – Management network
    • For RA administrators to login when patching the appliance
  – Ingest network
    • For receiving backups from protected databases or restoring backups from RA
  – Replication network
    • For replicating protected database backups between Recovery Appliances

• Bond Ingest & Replication for HA
  – Options are active/passive or 802.3ad
Architecture – Tape Options

• Two options for tape backups
  – Oracle integrated
    • Preconfigured Oracle Secure Backup (OSB)
    • Direct attach tape library via Recovery Appliance certified Fiber Channel Adapters installed in each compute server
    • Can connect to any OSB certified tape library
  – Third-party tape systems
    • Uses products like NetBackup
    • Backups sent via 10GigE to Media Manager
      – Normally using Ingest Network for transfer
    • Integration and operational support provided by 3rd Party Vendor
Program Agenda

1. Architecture
2. Backup Best Practices
3. Restore and Recovery
4. Recovery Appliance and Data Guard
5. Validation, Security and Troubleshooting
6. Best Practices at work
7. SK Hynix Recovery Appliance Use Case
Backup Manageability Best Practices

• Configure protected database
  – Use Enterprise Manager Cloud Control
    • Simplest deployment and configuration for 11g and 12c
  – Steps to backup database
    • Create Protection Policy on Recovery Appliance (RA)
    • Add Protected Database to RA
    • Configure Backup Settings for Protected Database
    • Schedule “Oracle-Suggested Recovery Appliance Backup”
Backup Best Practices

- Database backup script is simple
  - `$ rman target <target string> catalog <catalog string>
    backup
device type sbt
  cumulative incremental level 1
  filesperset 1
  section size 32g
database
  plus archivelog
  not backed up
  filesperset 32;`
RMAN best practices for the Recovery Appliance

• Initially allocate 2 RMAN channels per Database Node
  – Do not over-allocate RMAN channels - can result in worse performance
  – Take only one level 0 backup - ensure indexing of this backup has completed before taking another
    • select bp_key from rc_backup_piece where tag = '&tag' and backup_type = 'D' and virtual = 'NO';
  – Take daily cumulative incremental level 1 to reduce database recovery time

• Most common bottleneck is network or protected database’s I/O system

• Use Transparent Data Encryption (TDE) instead of RMAN encryption

• Use native database compression instead of RMAN compression

• Use change tracking file for all protected databases
Backup Best Practices for Bigfile Tablespaces

• Use Section Size 32 GB for bigfiles or large data files. Benefits level 0 backups for 11g and 12c databases

• Only 12c supports Section Size for Incremental Level 1 backups to parallelize across multiple RMAN channels
  – 11g RMAN commands ignore the Section Size clause if specified for Incremental Level 1 backups
  – backup device type sbt cumulative incremental level 1 filesperset 1 section size 32g database plus archivelog not backed up filesperset 32;
Program Agenda

1. Architecture
2. Backup Best Practices
3. Restore and Recovery
4. Recovery Appliance and Data Guard
5. Validation, Security and Troubleshooting
6. Best Practices at work
7. SK Hynix Recovery Appliance Use Case
Restore and Recovery Best Practices

How long can your production database be down for? What is your RTO requirements?

• When there is no validated disaster recovery plan, this might happen to you
  – Database Failure Occurs – must be restored from backup
    • Backup was not available on disk
    • Backup restored from tape
    • Found some tapes had been expired by mistake – took days to re-scan and re-catalog the pieces
    • Tape library had issues – moved tapes to another library that was only 1GigE connectivity
    • Tape restores were failing after many hours – init.ora parameter was wrong
  – Database was eventually handed back to customers but after 8 days and with significant data loss (RPO)
Restore and Recovery Best Practices

• Have a disaster recovery plan and rehearse your plan

• Use RMAN Restore Database / Recover Database as you would today
  – No new RMAN commands to learn
  – RMAN is aware of the Virtual Backups and will make the best decision for you
  – Can restore directly from tape or RA Replica without staging data on local RA

• Performance considerations
  – You need to think about parallelism and the impact on the protected database servers where the restore is going to
    • Too many channels might impact other databases in a consolidated environment
  – Will use the ingest network of the RA to restore the database
Restore and Recovery Best Practices

• Bigfile Tablespace Practices and Considerations
  – 11g and 12c supports Section Size for restore of level 0 backups
  – Only 12c supports Section Size for restoring a virtual backup from Recovery Appliance
    • Without Section Size support parallelism of bigfile data files and tablespaces are not possible
Program Agenda

1. Architecture
2. Backup Best Practices
3. Restore and Recovery
4. Recovery Appliance and Data Guard
5. Validation, Security and Troubleshooting
6. Best Practices at work
7. SK Hynix Recovery Appliance Use Case
Recovery Appliance and Data Guard

• Follow all MAA recommendations
  – Recovery Appliance per data center
  – Backup primary and standby databases to the local RA
  – No Recovery Appliance replication if standby already exists in the targeted data center
  – Restore operation can use any RA in any location
Recovery Appliance and Data Guard

MAA White Paper

• Post Data Guard role transition
  – No change in backup operations. Continue to backup both the primary and standby databases to the local RA

• Deploying the Zero Data Loss Recovery Appliance in a Data Guard Configuration
Program Agenda

1. Architecture
2. Backup Best Practices
3. Restore and Recovery
4. Recovery Appliance and Data Guard
5. Validation, Security and Troubleshooting
6. Best Practices at work
7. SK Hynix Recovery Appliance Use Case
Validation, Security and Troubleshooting

Top problems faced in the field

• RTO or RPO SLA’s not met
  – Bad Tapes
  – Corruptions in backups
  – Missing pieces (archive logs, data files or control files)
  – No automation or end to end understanding of restore and recover process

• Problem Avoidance:
  – Weekly RMAN crosschecks,
  – Weekly or monthly RMAN backup or restore validate
  – Monthly or Quarterly end to end restore and recovery validation testing and automation
Validation, Security and Troubleshooting

How Recovery Appliance addresses the issues

• Ingesting Backups
  – Validate data blocks as they are read from source database and sent to appliance

• Indexing Backups
  – Blocks received are validated, compressed as they are written to the delta store

• Ongoing Validation
  – All backupsets are crosschecked daily
  – All data file blocks are optimized weekly (meaning all blocks are read weekly)
  – All backupsets are validated (think restore validate) bi-weekly
  – May be modified by setting of configuration parameter
Validation, Security and Troubleshooting

How Recovery Appliance addresses the issues (continued)

• Built on Exadata
  – Benefits from Exadata Disk Scrubbing and Exadata Checksum checks

• All checks run on the Recovery Appliance offloading additional load on the protected databases

• Recovery plan still needed to be tested
  – Does not remove the need for periodic end to end restore and recovery testing to prepare operations team and validate issues outside RA
Validation, Security and Troubleshooting

Customers requiring end to end security

• Client to Recovery Appliance, or Recovery Appliance to Client
  – Database TDE is recommended
  – https, sqlnet encryption & Wallets/Certificates integration underway
  – MAA paper out soon

• Security in the Recovery Appliance
  – Recovery Appliance administrators responsibilities
    • Create Virtual Private Catalog (VPC) User
    • Assign protected databases to a specific VPC User
    • The protected database administrator can see all databases that share a common VPC user
Validation, Security and Troubleshooting

Troubleshooting Note

• For performance related issues on the Recovery Appliance refer to
  – Recovery Appliance Performance Issues Data Gathering Document (Doc ID 2066528.1)

• For network performance related issues between protected databases and the Recovery Appliance refer to
  – Recovery Appliance Network Test Throughput script (Doc ID 2022086.1)
Program Agenda

1. Architecture
2. Backup Best Practices
3. Recovery Appliance and Data Guard
4. Restore and Recovery
5. Validation, Security and Troubleshooting
6. Best Practices at work
7. SK Hynix Recovery Appliance Use Case
TEST CASE #1 – Complete Level 0 (Full Backup) within 24 hours

• Execute Level 0 backup on 200 protected databases while monitoring throughput and Recovery Appliance virtual full creation (backup indexing) activity

• **Passed:** Completed in 6 hours and 17 minutes, 4X faster than requirement. Backup rate was 14.7 TB/hour (4.2 GB/sec).

• Exceeded Customer Expectations. No Recovery Appliance tuning required.

<table>
<thead>
<tr>
<th>Task</th>
<th>Start</th>
<th>End</th>
<th>Duration (hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backup - L0</td>
<td>5/22/15 16:38</td>
<td>5/22/15 22:54</td>
<td>6.27</td>
</tr>
<tr>
<td>Indexing - L0</td>
<td>5/22/15 17:30</td>
<td>5/23/15 0:17</td>
<td>6.79</td>
</tr>
</tbody>
</table>
Test Case #2 – Copy 200 Database Backups to Tape in 7 days

- After the Level 0 backups, a workload generation script was executed to induce ~12% random block changes in each database.
- Upon completion, Level 1 (incremental) backups were taken on all databases.
- Copy to tape job templates on the Recovery Appliance were created, scheduled, and executed.
- Total of 6960 tape backup tasks executed.
- Average of 258 tasks/hour with throughput of 125MB/sec/tape drive.
- **Passed:** Backup to tape completed in 2 days and 3 hours, 3X faster than requirement, maximizing the 4 tape drives (tape throughput expected to increase with additional drives)
- **Exceeded Customer Expectations.** No Recovery Appliance tuning required.
TEST CASE #3 - Restore 2 databases with Concurrent L1 Backups

• Restore 2 databases while Level 1 backups are executing on 198 databases, with redo transport enabled on 159 databases.

• Level 1 backups must complete in 8 hours and restore operations must complete in 8 hours.

• Passed: All incremental backups completed in 2 hours, 4X faster than the requirement. Both databases were restored in 2 hours, 4X faster than the requirement, with a restore rate of 225 GB/hour. Note that restores in absence of concurrent backup workload could maximize ingest network bandwidth to achieve 12-14 TB/hour.

• Exceed Customer Expectations. Customer was more concerned about meeting backup windows and reasonable restore windows rather than achieving peak restore rates. No Recovery Appliance tuning was required.
Program Agenda

1. Architecture
2. Backup Best Practices
3. Recovery Appliance and Data Guard
4. Restore and Recovery
5. Validation, Security and Troubleshooting
6. Best Practices at work
7. SK Hynix Recovery Appliance Use Case
SK hynix is the global leader in producing semiconductor, such as DRAM and NAND flash and System IC including CMOS Image Sensors. Since pilot production of Korea’s first 16Kb SRAM in 1984, SK hynix consistently led the industry with smaller, faster and lower power semiconductor. As the second largest manufacturer of memory semiconductor, SK hynix is at the forefront of the IT industry.

Customer Profile
- Name: JeongRyun Park
- Company: SK hynix
- Role in Organization: IT Planning Team Leader / Information Technology Office

About Oracle in Our Organization
- Total 40 Unit of EXADATA across 4 business area in 5 FABs
- Hi-Tech Manufacturing Main System
- MES and Manufacturing Related System (Automation)
- HR, ERP, DW and Non MES System (Information)
Background of DR Systems

BCP (Business Continuity Plan) Introduction on Entire IT System.

• Sep 2013, SK hynix China Wuxi Factory Fire
  – Problems in production for 3 months

• Necessity of DR Systems
  – Manufacturing Automation Systems Campus DR
  – Information Systems Remote DR

SK Hynix Inc. Reports Fiscal Year 2013 and Fourth Quarter Results

Consolidated fourth quarter revenue was 3.4 trillion won decreased 18% from 4.1 trillion won of the previous quarter, due to decrease in production on account of Wuxi fab affected by a fire and the appreciation of Korean Won.
Benefit of Recovery Appliance

SK hynix ZDLRA Effect-> Quantitative Measurement

- Recovery Appliance reduce backup window to 12X
- Recovery Appliance reduce recovery window to 3X
- Recovery Appliance save space to 6X
Characteristics of Hi-Tech MES & Related Systems

**BigData & VLDB**
- Avg. DB Size: 30~50TB
- Daily Data Increment: 3~5TB
- Structure/Un-Structure
- Business Data / Sensor Data

**Extreme Transaction Processing**
- 4,000 ~ 6,500 Sessions
- 2,000 ~ 3,500 tx/sec
- Redo Gen. 80~100MB/Sec
- DW / Analysis Query
- Powered by Exadata

**Real-Time**
- Manufacturing Control
- Process Control
- Sensor Data Collection, Real-Time Analysis, Monitoring and Order Process
- No deferred processing

**Zero-Downtime**
- 24 * 7 * 365 No Down Time
- No Delay on Extreme Transaction
- MAA

Flexible Management on Big Data Not Impacting Real-Time Transaction
EXADATA has been services more than 40 systems in FDC and QA since 2012

- EQP Control Data
  - Line/Chamber Auto Control
  - MVIN/OUT, Remote Control
  - LOT Auto Reservation
- Line Interlock
  - Recipe/Reticle/DCOP/FDC
- DCOL
  - Production/Measuring Data (rSPC)

- FDC/SPC Data
  - Trace Data per Area
  - Line Tool Event per Area
  - Specific Line Sensor Data

- FDC Main Function and Practice: Real Time Data Collection, Fault Detection & Response Automation, Classification Function
- FSA / TAS Main Function and Practice: FDC and SPC Data Collection, QA , Reporting and Mining

- Production history
- Line history
- DCOL

- Overall Production History

- EQP
- EAP
- MES AP
- MES DB
- FDC AP
- FDC DB
- EDB / RPT
- FSA / TAS

- 5TB
- 8TB
- 19Units, 80TB
- 14Units, 70TB
- 19Units, 80TB
- 14Units, 70TB
EXADATA Legacy backup and Issues

Needs Fast Backup on Big Data Not Impacting Real-Time Transaction

AS-WAS Backup Configuration

- Each Exadata had been backed up to Legacy System
- FDC System
  - RMAN Backup
    - Image Copy
    - Incremental Update
  - Network
- EXADATA (45TB)
- Legacy System

Main Issue

- Impacting on Real-Time Transaction
  - Consumes CPU on Database Server and I/Os on Storage Server
  - Exadata resource management (DBRM and IORM) limit the backup resource usages under 40%

- Backup Window and Performance Impact Increasing
  - Full backup time on initial Half Rack was 7 hours → After Expansion to Full Rack, backup took 13 hours
  - 3 Hours for Incremental Update Backup

- Incremental Backup Validation Issue
  - Needs periodic backup validation
  - Validate Backup using Snapshot Clone

- Backup Management and Monitoring Overhead
  - Manage RMAN Scripts for 40 systems
  - Manage Scheduling and Monitoring for 40 systems.
Benefits of Recovery Appliance

EXADATA DR System Architecture

**Primary Center**

- **M14FDC**
  - Exadata X5-2 Half Rack
  - 7EF + 4 HC (HIGH)

- **M14LFDC**
  - Exadata X5-2 Half Rack
  - 7 EF (HIGH)

**DR Center**

- **ZDLRA**
  - Exadata X5-2 Half Rack

- **ADG**
  - M14 DR

- **EXADATA**
  - FDC
  - LFDC

- **Restore & Bring-up**

- **Flashback**

- **Incremental Forever**

**Benefits**:

- **Eliminate Impact on Real-Time Transaction**
  - Maximum Availability Architecture
  - Backup from Standby Database to minimize impact
  - Enable Flashback on Standby Database and Enable MAA parameters for comprehensive data protection

- **Simple, Low Overhead and Consistent Performance**
  - Only changed blocks are backed up → steady backup performance
  - Multiple Databases are covered by One ZDLRA
  - Guaranteed Restore Rate

- **Incremental Forever Strategy and Validation**
  - Backup window is very small consistently
  - Backup Validation is performed periodically

- **Management Benefits**
  - Easy backup environment configuration using Enterprise Manager
  - Archivelog Auto Backup
  - Do not need any backup configuration after Fail-Over
  - Intelligent Backup Space Estimation
### Recovery Scenario (RTO < 1 Hour)

<table>
<thead>
<tr>
<th>Target</th>
<th>Incident Type</th>
<th>Recovery Target</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ZDLRA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Recovery Time Actual</td>
</tr>
<tr>
<td>SPFILE LOSS</td>
<td>LOSS</td>
<td>&lt; 10 min</td>
</tr>
<tr>
<td>CONTROLFILE LOSS</td>
<td>LOSS</td>
<td>&lt; 10 min</td>
</tr>
<tr>
<td>BLOCK Corruption</td>
<td></td>
<td>&lt; 10 min</td>
</tr>
<tr>
<td>REDO LOG</td>
<td>Current Redo Loss</td>
<td>&lt; 10 min</td>
</tr>
<tr>
<td>REDO LOG</td>
<td>Active/Inactive Redo Loss</td>
<td>&lt; 10 min</td>
</tr>
<tr>
<td>TABLE/PARTITION</td>
<td>TABLE Loss</td>
<td>&lt; 30 min</td>
</tr>
<tr>
<td>TABLE/PARTITION</td>
<td>PARTITION Loss</td>
<td>&lt; 30 min</td>
</tr>
<tr>
<td>DATAFILE</td>
<td>Specific Datafile Loss</td>
<td>&lt; 15 min</td>
</tr>
<tr>
<td>TABLESPACE</td>
<td>Specific Tablespace Loss</td>
<td>&lt; 30 min</td>
</tr>
<tr>
<td>TABLESPACE</td>
<td>System Tablespace Loss</td>
<td>&lt; 30 min</td>
</tr>
<tr>
<td>TABLESPACE</td>
<td>Undo Tablespace Loss</td>
<td>&lt; 30 min</td>
</tr>
<tr>
<td>DATABASE</td>
<td>Fresh DB Creation + TPITR</td>
<td>&lt; 60 min</td>
</tr>
<tr>
<td>SITE FAILURE</td>
<td>Site Failure</td>
<td></td>
</tr>
</tbody>
</table>

1) Among 40 TB, Recent 7 days of data should be recovered first for service open. Rest of data will be recovered after service open.