

Backup and Recovery
Performance and Best Practices
for Exadata Cell and Oracle
Exadata Database Machine

Oracle Database Release 11.2.0.2 and 11.2.0.3

*Oracle Maximum Availability Architecture White Paper
February 2013*

Maximum Availability Architecture

Oracle Best Practices For High Availability

Executive Overview	1
Introduction	2
Key Performance Observations and Rates.....	4
Performance Results for Disk-Based Backup and Restore	5
Performance Results for Tape-Based Backup and Restore.....	6
Estimating Tape Backup Throughput Rates.....	7
Media Management Software for Tape Backups	9
Tape-Based Backup Strategy.....	10
Best Practices for Tape-Based Configurations	13
Database Configuration Best Practices	14
RMAN Configuration Commands and Backup Scripts	18
Configuring InfiniBand Network to Media Server	18
Configuring the Gigabit Ethernet (GigE or 10GigE) Network to Media Server	22
Configuring Persistent Bindings for Tape Devices	23
Backing up the Oracle Secure Backup Catalog	24
Disk-Based Backup Strategy	24
Best Practices for Disk-Based Configurations.....	26
Database Configuration Best Practices	26
RMAN Configuration Commands and Backup Script	28
Restore and Recovery Best Practices	29
Offload Backups with Oracle Data Guard	30
Caution when Considering non-Exadata Storage for Backups.....	30
Monitoring and Troubleshooting	31
Monitoring RMAN	31
Monitoring and Troubleshooting Oracle Secure Backup	31
Monitoring TCP/IP Traffic	32

Conclusion 33

Executive Overview

The integration of Oracle Exadata Database Machine and Oracle Maximum Availability Architecture best practices (Exadata MAA) provides the most comprehensive HA solution available for Oracle Database.

A key operational aspect of deploying Exadata Database Machine is to be sure that database backups are performed and Oracle Database can be restored if disaster strikes. This white paper is based on Oracle Database release 11.2.0.2 or higher releases and describes the best practices for setting up the optimal backup and recovery strategy to protect mission-critical data. For Exadata Database Machine backup and recovery best practices applicable to Oracle Database release 11.2.0.1, see the MAA white paper at <http://www.oracle.com/technetwork/database/features/availability/maa-tech-wp-sundbm-backup-final-129256.pdf>.

This paper discusses the following topics:

- Key performance observations and rates
- Recommended tape-based backup solution
- Tape-based configuration best practices
- Recommended disk-based backup solution
- Disk-based configuration best practices
- Restore and recovery best practices

The January 2013 update of this paper includes performance rates using Exadata Storage Expansion Rack and using High Capacity Disk Drives throughout. The testing was performed with Oracle Database release 11.2.0.3.

Introduction

Oracle Database has very sophisticated and scalable backup technologies. These technologies work especially well on Exadata Database Machine with its high bandwidth InfiniBand network and Exadata Storage Server Grid. Combined these Oracle technologies provide database-aware storage services, such as the ability to offload database backup processing from the database server to storage, while remaining transparent to SQL processing and database applications. A case in point is the Exadata Database Machine X2-2 or X3-2 Full Rack configuration described in this paper, for which tape-based backups achieved over 8.6 TB/hour for full backups and an effective backup rate of 10 to 50 TB/hour for incremental backups. Higher full backup rates are possible by using more media servers and tape drives than configured in the tests described in this paper.

For disk-based backups and with Oracle Database release 11.2.0.3 and higher, full image copies achieved over 20 TB/hour and an effective backup rate of 20 to 50 TB/hour was attained for incremental backups to the local Exadata Storage.

Database restore rates achieved over 20 TB/hour into normal redundancy disk group.

Disk-based backup and restore using a full Exadata Storage Expansion Rack (18 Exadata cells) achieved rates of 27 TB/hr for both database backups and database restores.

Database redo apply is able to achieve 2.7 TB/hour (800 MB/sec) for batch or 700 GB/hour (200 MB/sec) for intensive OLTP workload when using Exadata write-back flash cache. .

The above rates are achieved with database CPU utilization of less than 5% on the targeted database servers, leaving plenty of CPU bandwidth for concurrent user workloads.

To put these backup, restore, and redo apply rates in perspective, a 30 TB database which consumes 100 TB of data when mirroring and temp are accounted for, can be backed up in less than 2 hours when performing a full backup or in less than 1 hour for an incremental backup. Restoring the same complete backup can take less than 2 hours and redo apply can recover 2 TB of redo in an hour. On Exadata Database Machine the factor that limits tape backup and restore rates is typically hardware external to the Database Machine, not the Database Machine or the Oracle backup and recovery software.

The following technologies make it possible to attain these backup rates:

- Exadata Database Machine is an integrated system of software, servers, storage, and networking for all data management, including data warehousing, transaction processing, and consolidated mixed application workloads:

- The InfiniBand fabric provides an extremely high performance network for transferring backup data from storage servers to database servers and then to tape media servers, and vice versa.
- While transfer rates are high, the overhead on CPU resources used for InfiniBand network transfers is very low compared to other solutions
- The Exadata Storage Server runs the Exadata Storage Server Software provided by Oracle. The hardware components of the Exadata Storage Server (also referred to as an Exadata cell) are carefully chosen to match the needs of high performance database processing. Exadata software is optimized to take the best possible advantage of all Exadata hardware components and Oracle Database. Each Exadata cell delivers outstanding I/O performance and bandwidth to the database.
- Exadata Database Machine is optimized for use with Oracle Database applications. It delivers very high I/O and SQL processing performance for online transaction processing (OLTP), data warehousing (DW), and a consolidation of mixed workloads.
- Exadata Storage Server offload capability combined with RMAN block change tracking will efficiently perform large I/Os at the storage-tier level, returning only individual changed blocks for incremental backups and increasing the backup performance of the system.
- Exadata Hybrid Columnar Compression (EHCC) and native Oracle Database compression capabilities can significantly reduce the overall database size and the I/O.
 - Oracle OLTP table compression can provide compression rates with a factor of 2x to 4x, such that a table normally requiring 100 GB of disk space would require only 25 GB to 50 GB of disk space.
 - Exadata Hybrid Columnar Compression (EHCC) can provide compression rates with a factor of 10x to 50x. Thus, the same table that requires 100 GB disk space could require only 2 GB to 10 GB of disk space after compression, depending on the data. This means that user data backup and restore rates can be an order of magnitude higher than the physical rates described previously. For example, after a 10x EHCC compression of user data is taken into account, the effective backup rate increases from 7 TB/hour of physical data to 70 TB/hour of user data.
- Oracle Recovery Manager (RMAN) provides the native backup and recovery infrastructure within Oracle Database, enabling optimized data protection in Exadata environments:
 - Backup, restore, and recover operations are performed using standard RMAN commands.
 - RMAN backup and recovery combined with Exadata Hardware Assisted Resilient Data (HARD) provides a comprehensive, data aware solution to:
 - Prevent block corruptions from being written to disk.

- Allow automatic and manual repair options using Oracle ASM automatic repair mechanisms, Exadata storage built-in corruption checks, and the Oracle generic block corruption detection parameters.
- RMAN can parallelize backup operations across all database nodes and Exadata storage cells. This allows all disks, all network connections, and all CPUs in the system to contribute to performing backup operations.
- RMAN block change tracking allows incremental backups to run very quickly and efficiently. With block change tracking, only the areas of the database that have been modified since the last incremental backup or full backup are read from disk.
- Oracle Secure Backup is a centralized tape backup management solution for the entire IT environment including file system and Oracle Database. With built-in RMAN integration, Oracle Secure Backup delivers the fastest Oracle Database backup to tape. Some important backup optimizations that provide substantial savings in backup time and tape costs are available only with Oracle Secure Backup and RMAN:
 - Unused block compression eliminates the time and space usage needed to backup blocks that are allocated to tablespaces but are not currently used by tables.
 - Undo optimization eliminates the time and space usage needed to back up undo that is not required to recover the current backup.

Key Performance Observations and Rates

This paper describes the backup and recovery performance testing of Exadata Database Machine X2-2 Full Rack, Half Rack, and Quarter Rack and Exadata Database Machine X2-8 across various configurations:

- Disk backup and restore testing was performed with image copy formats using a fast recovery area located on Exadata storage and using varying degrees of RMAN parallelism.
- Tape backup and restore testing was performed using Oracle Secure Backup Release 10.4, with 2 media servers attached to a Sun StorageTek SL3000 tape library and 8 Sun StorageTek T10000C tape drives via a Storage Area Network (SAN).

The tables in the following sections summarize performance results for disk-based and tape-based backup and restore performance.

Performance Results for Disk-Based Backup and Restore

TABLE 1 : SUMMARY OF MEASURED DISK-BASED BACKUP AND RESTORE PERFORMANCE

FULL DATABASE BACKUP TO DISK USING IMAGE COPIES ¹²			
Instances and Channels	Quarter Rack	Half Rack	Full Rack
X2-2 or X3-2 (11.2.0.3) All instances 2 RMAN channels per instance	N/A	N/A	20-25 TB/hour (RECO is NORMAL redundancy disk group) 27 TB/hour (BKUP is NORMAL redundancy disk group on Storage Expansion Rack)
X2-8 or X3-8 2 instances 8 RMAN channels per instance	N/A	N/A	20-25 TB/hour (RECO is normal redundancy disk group) 27 TB/hour (BKUP is NORMAL redundancy disk group on Storage Expansion Rack)
X2-2 (11.2.0.2) All instances 2 RMAN channels per instance	4 TB/hour	8-9 TB/hour	17-18 TB/hour
X2-8 (11.2.0.2) 2 instances 8 RMAN channels per instance	N/A	N/A	17-18 TB/hour
X2-2 (11.2.0.1) 2 instances, 4 RMAN channels per instance	1589 MB/sec or 5.4 TB/hr	2081 MB/sec or 7.1 TB/hr	2081 MB/sec or 7.1 TB/hr

¹ Note that a full database backup to disk using image copies only occurs once, because subsequent disk backups are performed using RMAN incremental backup and merge.

² Note that a patch for Bug# 13355365 is required to get the X2-8 or X3-8 high backup and restore rates

FULL DATABASE INCREMENTAL BACKUP TO DISK (10% CHANGE) USED DAILY

2 instances, 2 RMAN channels Measured effective backup rate 20 to 50 TB/hr, depending on workload per instance

FULL DATABASE RESTORE FROM DISK

X2-2 or X3-2 (11.2.0.3)	N/A	N/A	20-25 TB/hour (DATA is NORMAL redundancy disk group)
X2-8 or X3-8 (11.2.0.3)	N/A	N/A	20-25 TB/hour (DATA is NORMAL redundancy disk group)
X2-2 (11.2.0.2) All instances 2 RMAN channels per instance	4 TB/hour	7 TB/hour	14TB/hour (DATA is HIGH redundancy disk group) 17 TB/hour (DATA is NORMAL redundancy disk group)
X2-8 (11.2.0.2) 2 instances 8 RMAN channels per instance	N/A	N/A	14 TB/hour (DATA is HIGH redundancy disk group) 17 TB/hour (DATA is NORMAL redundancy disk group)

For optimal performance of disk-based backups, incremental backups and database restores on Exadata Database Machine X2-2 use all database instances and two or four RMAN channels per instance. For optimal performance on Exadata Database Machine X2-8 use all database instances and up to 8 RMAN channels per instance. Less than 5% CPU was used in either configuration.

Performance Results for Tape-Based Backup and Restore

TABLE 2 : SUMMARY OF TAPE-BASED BACKUP AND RESTORE PERFORMANCE

FULL DATABASE BACKUP TO TAPE			
Instances and Tape Drives	Quarter Rack	Half Rack	Full Rack
All instances, 8 tape drives (T1000C). 1 RMAN channel per tape drive	N/A	N/A	2621MB/sec or 9 TB/hour or 327MB/sec per tape drive
All instances, 14 tape drives	2509 MB/sec or 8.6 TB/hr	2509 MB/sec or 8.6 TB/hr	2509 MB/sec or 8.6 TB/hr

(LTO4), 1 RMAN channel per tape drive	or 179 MB/sec per tape drive	or 179 MB/sec per tape drive	or 179 MB/sec per tape drive
FULL DATABASE INCREMENTAL BACKUP TO TAPE (10% CHANGE)			
All instances, 14 tape drives (LTO4), 1 RMAN channel per tape drive	Measured effective backup rate 10 to 70 TB/hr		
FULL DATABASE RESTORE FROM TAPE			
All instances, 8 tape drives (T1000C). 1 RMAN channel per tape drive	N/A	N/A	2300 MB/sec or 8 TB/hour or 291 MB/sec per tape drive
All instances, 14 tape drives (LTO4), 1 RMAN channel per tape drive ³	1800 MB/sec or 6.1 TB/hr or 128 MB/sec per tape drive	2271 MB/sec or 7.8 TB/hr or 162 MB/sec per tape drive	2271 MB/sec or 7.8 TB/hr or 162 MB/sec per tape drive

The limiting factor for tape-based backup and restore performance described in Table 2 was the number of tape drives used (except in the case of a Quarter Rack tape restore where performance is limited by the number of available disks). For all tape-based backup and restore operations into existing files, CPU utilization on the database nodes was less than 10% and on the media server nodes it was less than 20%. Best practices also recommend utilizing all available database instances for tape-based backup and restore operations.

Estimating Tape Backup Throughput Rates

Answer the following questions and then use the worksheet examples shown in Table 3 to quickly evaluate potential architectural bottlenecks and determine the optimal tape backup architecture for your requirements.

Questions:

³ The restore rate achieved for Exadata Database Machine Full Rack is also applicable for Exadata Database Machine Half Rack. However, this rate should not be expected for a Exadata Database Machine Quarter Rack since restore performance is limited by the number of disks.

- What is the slowest network segment between Exadata Database Machine and the tape libraries?
- What is the media server ingestion or transfer rate? (Media Servers are described in [“Media Management Software for Tape Backup”](#))
- What is the tape drive transfer rate?
- How many tape drives do you plan to use?
- What is the expected tape compression ratio?

Table 3 shows tape-based backup flow and highlights the throughput of each component:

- The content shown in **bold** typeface depicts example data. However, you can use the values in the Estimated Rate column for each backup estimate.
- The Throughput Rate (in the last column of Table 3) is determined using the following equation:

$$\text{Quantity} \times \text{Estimated Rate} = \text{Throughput Rate}$$

TABLE 3 : EXAMPLE TAPE BACKUP WORKSHEET

FLOW	COMPONENTS	QUANTITY	ESTIMATED RATE (GB/SEC)	THROUGHPUT RATE (GB/SEC)
1	Exadata Cell	14	1 ⁴	14
2	Database Server	8	2.0 ⁵	16

⁴ SAS-based Exadata Cell supports up to 1.5 GB/sec data bandwidth. Measured values were used for the estimates shown in Table 3.

3	Network to Media Server:			
	a) Media Server InfiniBand HCA using TCP/IP	2	2	4
	b) Media Server 10GigE Active/Active NICS	4	1	4
	c) Media Server GigE NICS	4	0.12	0.48
4	Media Server to Tape Library SAN ⁶ Links	4	0.8	3.2
5	Tape Drives (LTO4)	14	0.17	2.3

Use 1 GB/sec for SAS or 700 MB/sec for SATA-based Exadata Cell

For example, using the throughput rates shown in the last column of Table 3, the following bottlenecks will be reached first if using InfiniBand fabric to media server:

1. Media server to tape library SAN transfer rate
2. Number of tape drives

To increase the backup rate, add more host bus adaptors (HBAs) to the media server or tape library, and increase the number of tape drives.

If using Gigabit Ethernet (GigE) network between the media server and Exadata Database Machine as shown in the row labeled “3c) Media Server GigE NICS” in Table 3, then the Media Server GigE NICs throughput rate will be your main bottleneck.

Media Management Software for Tape Backups

RMAN optimizes Oracle Database backups to disk and RMAN is integrated with media management software for backup to tape. Media management software⁷ is the software layer that facilitates RMAN backups to tape.

⁵ Each server has Dual-ported Quad Data Rate (QDR) InfiniBand Host Channel Adapters (HCAs) (40 Gb/s). However, while performing backups over InfiniBand using the TCP/IP protocol, the achievable throughput is approximately 2 GB/sec.

⁶ Storage Area Network (SAN).

Oracle Secure Backup is the media management software used for the testing described in this white paper. It is highly scalable with a client/server architecture in which all hosts in the backup domain have one or more Oracle Secure Backup roles. The roles include:

- **Administrative server** is the host that stores configuration information and catalog files for hosts in the administrative domain. A backup domain includes only one administrative server. One administrative server can service every client on the network. The administrative server runs the scheduler which starts and monitors backups within the administrative domain.
- **Media servers** are computers or servers to which at least one tape device is connected. A media server is responsible for transferring data to or from the tape devices that are either directly attached or Storage Area Network (SAN) attached.
- **Client** is any computer or server whose files Oracle Secure Backup backs up or restores. All computers to be backed up within the domain will be assigned the client role during installation along with additional roles as defined by the user, including the media server and/or administrative server roles.

Any computer in the domain can act as the administrative server. In practice, it is common to have an additional computer host the Oracle Secure Backup catalog, RMAN catalog, and the Oracle Enterprise Manager repository.

Backing up and restoring file system data is not addressed in this paper beyond protecting the Oracle Secure Backup catalog. For more information about configuring file system backup and restore, see the Oracle Secure Backup documentation at http://download.oracle.com/docs/cd/E14812_01/doc/doc.103/e12834/toc.htm

Tape-Based Backup Strategy

Some of the key benefits to tape-based backup strategies include:

- Exadata storage and tape-based backups provide extremely fast backup and restore rates.
- Tape-only solutions isolate faults from Exadata Storage Servers.
- Exadata Database Machine capacity and bandwidth are maximized.

⁷ The media management software described in this paper is Oracle Secure Backup. However, the configuration best practices are applicable to other media management software products.

However, disk-based backup may be a better solution if you require faster recovery times for data and logical corruptions and for some Tablespace Point in Time (TSPITR) scenarios. For more information about disk-based backup solutions, see the “[Disk-Based Backup Strategy](#)” section in this white paper.

For tape-based backup solutions, the recommended strategy is to perform the following backups:

- Weekly RMAN level 0 (full) backups of the database
- Daily cumulative RMAN incremental level 1 backups of the database
- Daily backups of the Oracle Secure Backup catalog

Use the following strategies to scale backup rates using the InfiniBand fabric while maintaining high availability:

1. Start with at least two media servers. There can be a total of two Host Channel Adapters (HCA) - only one active port per media server - bonded for high availability.
2. Add tape drives until the HCA of each media server’s bandwidth is saturated.
3. Use all Oracle Database server instances for backups.
4. Use one RMAN channel per tape drive.

If you are using a Gigabit Ethernet (GigE) network between the media server and Exadata Database Machine, your rate will be bound by the number of GigE connections multiplied by 120 MB/sec. In an Exadata Database Machine Full Rack there are eight HA-bonded GigE ports available or one HA-bonded GigE port per database node.

Figure 1 shows the Oracle MAA recommended architecture.

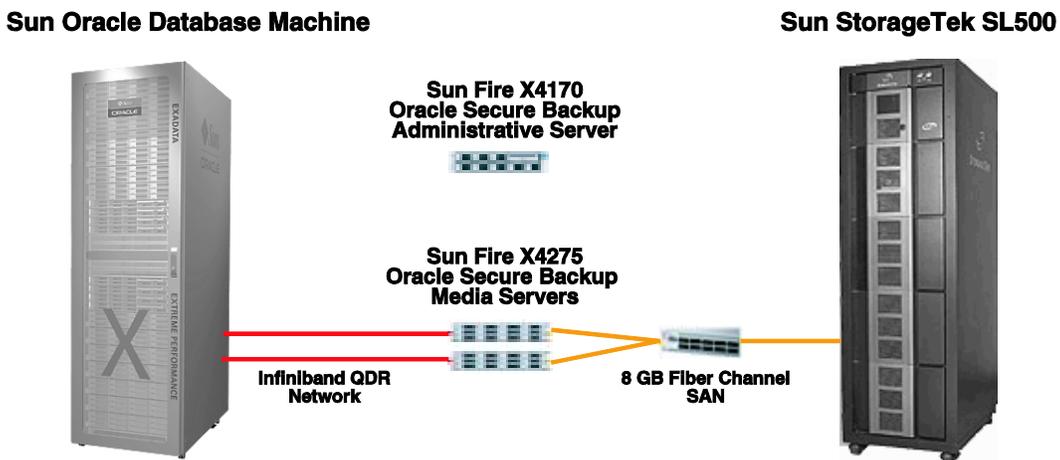


Figure 1 Exadata Database Machine Hardware System Configuration

Figure 1 describes a tape backup solution with two or more media servers. Connect media servers directly to the existing InfiniBand fabric using the six HA-bonded ports available for Exadata Database Machine. For example, use ports 5B, 6A, 6B, 7A, 7B, and 12A for any size (Full Rack, Half, Rack, or Quarter Rack) Exadata Database Machine. Any other available ports should be reserved in the event that you want to upgrade Exadata Database Machine Half Rack to a Full Rack.

Each media server requires an InfiniBand QDR HCA or the recommended dual-ported InfiniBand QDR HCA. The network protocol used for backups over InfiniBand is the standard TCP/IP protocol, so it is transparent to the backup software on the database servers and the media servers. The backup software operates identically whether using InfiniBand or a Gigabit Ethernet network.

Sun Oracle Database Machine

Sun StorageTek SL500

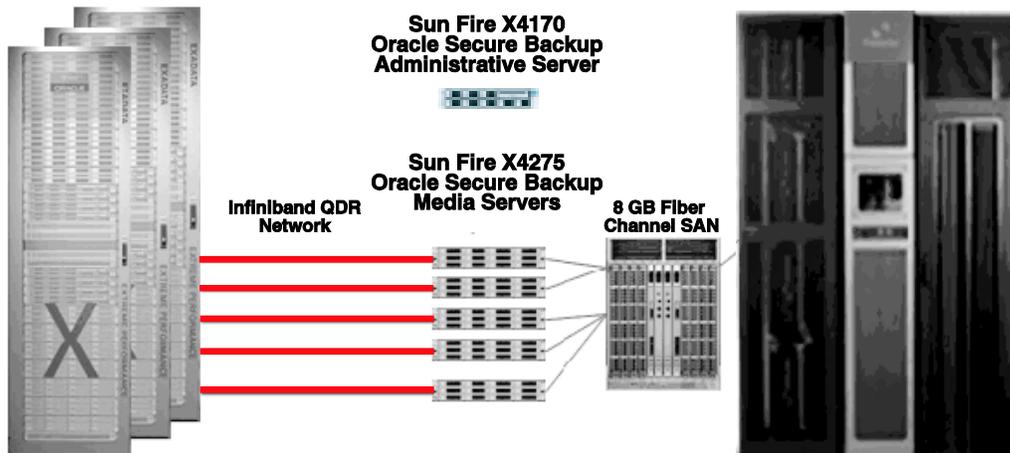


Figure 2: Larger Backup Configuration with Multiple Media Servers

Figure 2 describes a larger backup configuration having an effective maximum data transfer rate of 10 GB/sec, or 35 TB/hour. Allocate a sufficient number of tape drives to enable media servers to attain maximum backup and restore rates. For example, if a tape drive backup rate is

240 MB/sec of compressed data⁸, then use at least 8 tape drives to achieve the maximum data transfer rate of one media server's HCA.

Figure 3 describes the recommended Exadata Storage Server layout for tape-based backup and recovery. In the figure, the faster (outer) 80% of the disk is assigned to the DATA area, and the slower (inner) 20% of the disk is assigned to the fast recovery area (RECO) area. This layout can be configured automatically during deployment. Furthermore, for a tape-based only backup solution, MAA best practices recommend configuring a high redundancy RECO disk group containing OCR, Voting, spfiles, redo log groups, control files, and archives if high redundancy is not possible for the DATA disk group due to space constraints. Please refer to the following MAA best practices technical whitepaper for ASM configuration best practices on Exadata Database Machine:

<http://www.oracle.com/technetwork/database/features/availability/exadata-maa-131903.pdf>.

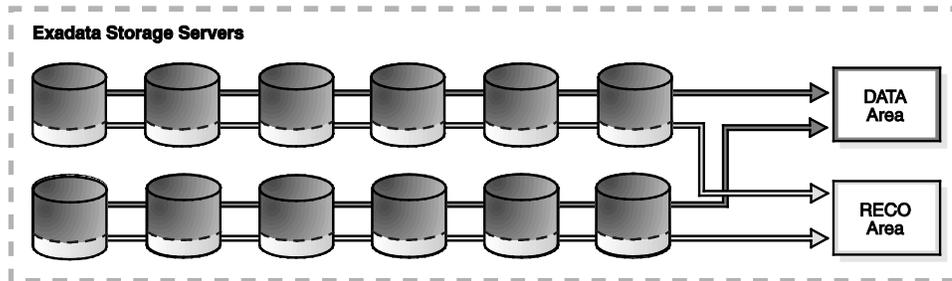


Figure 3: Exadata Storage Server Grid Disk Layout for Tape Based Backup and Recovery

Best Practices for Tape-Based Configurations

This section provides:

- [Database configuration practices](#)
- [RMAN commands and configuration](#)
- [Configuring InfiniBand Network to Media Server](#)
- [Configuring the Gigabit Ethernet \(GigE\) Network to Media Server](#)
- [Configuring persistent bindings for the tape drives](#)

⁸ An LTO4 Tape Drive is capable of writing approximately 240 MB/s compressed data to tape, while an LTO3 Tape Drive is capable of writing approximately 160 MB/s of compressed data to tape.

- [Backing up the Oracle Secure Backup catalog](#)

Database Configuration Best Practices

- **Use Oracle Secure Backup for low-cost, fast, and MAA-validated tape backups.**
 - Oracle Secure Backup provides the fastest database backup to tape due to its tight integration with RMAN.
 - If you are backing up to tape using Oracle Secure Backup, then the unused-block optimization capability is enabled. However, if the backup is made directly to tape using a third-party media management product, then this does not have any effect because the unused-block optimization is available only with Oracle Secure Backup.
- **Configure Network Time Protocol (NTP) daemon on Oracle Secure Backup servers.**
 - Ensure that the NTP daemon service on the Oracle Secure Backup Admin and Media Servers are running and configured to use the same time source as the Exadata Database Machine. Oracle Secure Backup schedules jobs based on the time of the database node, but the scheduler executes the job based on the time on the Oracle Secure Backup Admin server. If the time on the two systems is different, then a job can be queued on the admin server until the appropriate time.
- **Configure the Preferred Network Interface (PNI) to direct the Oracle Secure Backup traffic over the InfiniBand network interface.**

```
ob> lspni (List Preferred Network Interface)
mediaserver1:
  PNI 1:
    interface:  mediaserver1-ib
    clients:    dbnode1, dbnode2, dbnode3, dbnode4,
dbnode5, dbnode6, dbnode7, dbnode8
  PNI 2:
    interface:  mediaserver1
    clients:    adminserver
dbnode1:
  PNI 1:
    interface: dbnode1-ib
    clients:  mediaserver1
```

- **Configure one RMAN channel per tape drive and add tape drives to scale backup rates.**

Typical tape drive backup rates are between 100 MB/sec and 240 MB/sec, depending on the drive type and compression options. Note that tape drive compression becomes less effective when backing up tables that are compressed at the database level. Backup performance scales

when you add more tape drives and RMAN channels, assuming there is available throughput on the media server.

A single RMAN channel can stream data at a rate of 749 MB/sec from Exadata Database Machine to the media server. Backup performance, therefore, is limited by tape drive throughput.

- **Configure Gigabit Ethernet or InfiniBand.**

- Use Gigabit Ethernet for smaller databases:
 - Using a dedicated interface for the transport eliminates the impact to the client access network. Typically, a dedicated backup network is in place. The maximum throughput with the Gigabit Ethernet network is 120 MB/sec multiplied by the *<Number of Database Servers>*.
 - A Full Rack Exadata Database Machine can achieve throughput up to 960 MB/sec.
 - For more information, see the [“Configuring the Gigabit Ethernet \(GigE or 10GigE\) Network to Media Server”](#) section in this white paper.
- Use 10 Gigabit Ethernet for larger databases:
 - Using a dedicated interface for the transport eliminates the impact to the client access network. Typically, a dedicated backup network is in place. The maximum throughput with the 10 Gigabit Ethernet network is 1 GB/sec times the *<Number of Database Servers>*.
 - For a Full Rack Exadata Database Machine, achieving throughput as high as 8 GB/sec is possible.
 - For more information, see the [“Configuring the Gigabit Ethernet \(GigE or 10GigE\) Network to Media Server”](#) section in this white paper.
-
- Use InfiniBand for the best backup rates, especially for larger databases that require fast backup rates and low CPU overhead.

The graph in Figure 4 shows the CPU Utilization from one of the database nodes involved in the backup using Oracle Secure Backup when running over InfiniBand Network.

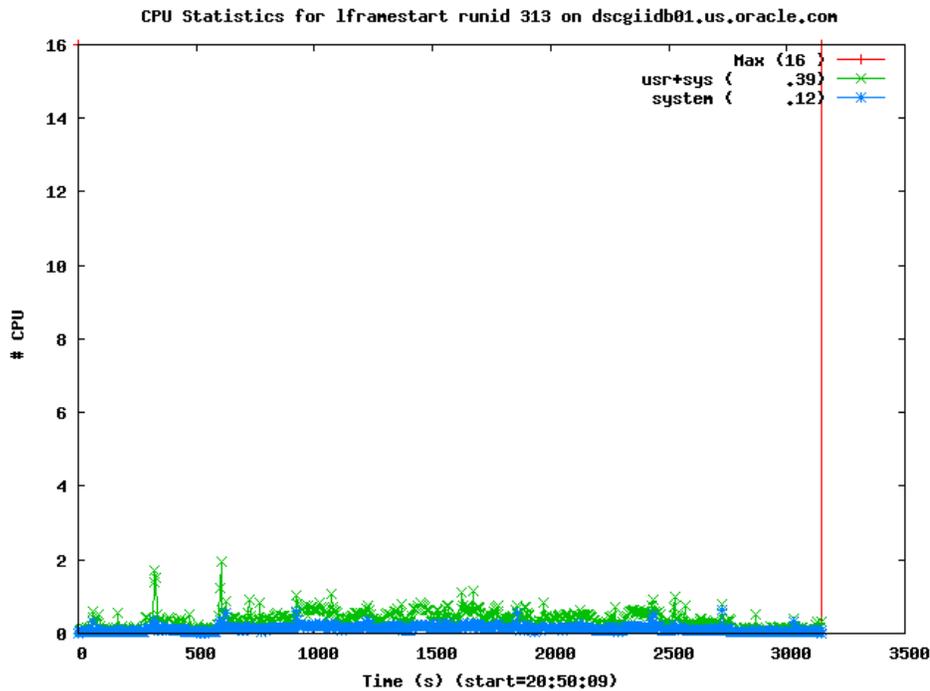


Figure 4: Exadata Storage Server Grid Disk Layout for Tape Based Backup and Recovery

For more information, see the “[Configuring InfiniBand Network to Media Server](#)” section in this white paper.

- **Configure Oracle RAC Service for backup running on all database instances.**

CPU utilization on the database and media server nodes was less than 10% when spread across all database instances.

- **Use SQL*Net service load balancing to distribute RMAN channels evenly among the allocated instances.**

- 1) Create a service that runs on the selected nodes in the cluster:

```
srvctl add service -d <dbname> -s <service name> -r
<instance1>,<instancen>
```

- 2) When running RMAN, use the service name in the connect string for the “target” parameter:

```
rman target sys/<password>@<scan_address>/<service_name> catalog ...
```

Note that in order to get true load balancing, refer to Appendix-A

- **Use RMAN incremental backups and block change tracking.**

Enable block change tracking to achieve fast incremental backups. Block change tracking allows RMAN to avoid scanning blocks that have not changed when creating incremental backups. Also, when performing incremental backups of databases on Exadata Cell, additional block inspection is offloaded from the database servers.

Block change tracking provides the greatest benefit for databases where fewer than 20% of the blocks are changed daily. You may still benefit by using block change tracking with change rates greater than 20% but testing is recommended to ensure that backup times are reduced.

- **Set `DB_RECOVERY_FILE_DEST_SIZE` to bound space in the fast recovery area.**

The database writes archived redo log files and any additional recovery files to the fast recovery area. These include any disk backup files such as level 0 image copies and level 1 backup sets as well as Flashback log files (if Flashback Database is enabled). It is important that the value of the `DB_RECOVERY_FILE_DEST_SIZE` parameter be set to less than the total free space in the disk group and the free space should take account of at least one disk failure and preferably one Exadata Cell failure.

Additionally, if multiple databases are sharing the fast recovery area, you must ensure that the sum of the space allocated to the different databases is less than the free space in the disk group.

- **Use an external RMAN recovery catalog.**

See the [Oracle Database Backup and Recovery User's Guide](#) for more information about the RMAN repository.

- **Consider using I/O Resource Management or Database Resource Manager to manage system resources on the Exadata Database Machine.**

If Exadata Database Machine's CPU or I/O resources must be prioritized between application workload and backups, then use I/O Resource Manager or Database Resource Manager. See the *Oracle Exadata Storage Server Software User's Guide* for instructions about setting up and configuring I/O Resource Manager.

- **Tune the network communication if you are using a third-party media management vendor.**

If you are using a non Oracle media management vendor, then contact the vendor for their configuration best practices. Most vendors test and validate their own products with the Exadata Database Machine and can recommend how to exploit the full potential of the InfiniBand or GigE networks. Note that there is no special certification required for RMAN and the Media Management Vendor (MMV) to work with Exadata. The MMV is required to certify with Oracle Database 11g release 2 (11.2) and Oracle Enterprise Linux.

RMAN Configuration Commands and Backup Scripts

- Use the following configuration commands to parallelize backups across all database nodes, allowing all the disks, network connections, and system CPUs to be leveraged for increased performance.

RMAN configuration changes:

```
CONFIGURE DEFAULT DEVICE TYPE TO SBT;  
CONFIGURE DEVICE TYPE SBT PARALLELISM <# Of Tape Drives>;
```

- Use the following backup scripts to automate weekly and daily backups.

RMAN script for weekly backup:

```
run {  
  backup incremental level 0 database;  
  backup archivelog all not backed up;  
}
```

RMAN script for daily backup:

```
run {  
  backup cumulative incremental level 1 database;  
  backup archivelog all not backed up;  
}
```

Configuring InfiniBand Network to Media Server

With the available InfiniBand ports in a Exadata Database Machine, media servers can be directly connected to the InfiniBand fabric by adding an InfiniBand Quad Data Rate (QDR) host channel adapter (QDR HCA) to the media server. For high availability, connect the HCA to two different Exadata Database Machine InfiniBand switches to eliminate the switch as a single point of failure. This provides transparent failover if connectivity is lost to one of the ports.

Follow these best practices.

Note: The examples included with each bullet are based on a media server running on the Linux operating system.

- **Configure bonding of the InfiniBand interfaces on the media server.**

The following example of bonding `ib0` and `ib1` is specific to a Linux environment:

1. Modify the `/etc/modprobe.conf` file to add the following two lines to the bottom of the file. This adds another bonding alias and options.

```
alias bondib0 bonding  
options bonding max_bonds=2
```

The file will be similar to the following example. This file assumes bonding was previously established on bond0:

```
alias eth0 tg3
alias scsi_hostadapter cciss
alias scsi_hostadapter1 ata_piix
alias scsi_hostadapter2 usb-storage
alias ib0 ib_ipoib
alias ib1 ib_ipoib
alias bondeth0 bonding
alias bondib0 bonding
options bonding max_bonds=2
```

2. Create the `/etc/sysconfig/network-scripts/ifcfg-bondib0` file, as follows.

```
DEVICE=bondib0
USERCTL=no
BOOTPROTO=none
ONBOOT=yes
IPADDR=<IP Address for bondib0 within the same subnet as the
existing InfiniBand network >
NETMASK=<Netmask must be the same as the existing InfiniBand
network >
NETWORK=<Network calculated using ipcalc-n
ip_address netmask>
GATEWAY=<Gateway IP address>
BONDING_OPTS="mode=active-backup miimon=100
downdelay=5000 updelay=5000"
IPV6INIT=no
```

3. Make copies of the current `ib0` and `ib1` configuration files. Ensure the copied files do not start with `ifcfg-ib0`. Prefix the file name with `backup-` or a similar word, and do not add a suffix such as `-backup`. For example:

```
cd /etc/sysconfig/network-scripts/
cp ifcfg-ib0 backup-ifcfg-ib0
cp ifcfg-ib1 backup-ifcfg-ib1
```

4. Modify the current `ib0` and `ib1` configuration files so they are configured to act as slaves to the `bondib0` interface. The files should appear as follows:

```
* File ifcfg-ib0:
DEVICE=ib0
USERCTL=no
ONBOOT=yes
MASTER=bondib0
```

```
SLAVE=yes
HOTPLUG=no
BOOTPROTO=none
MTU=65520
```

```
* File ifcfg-ib1:
DEVICE=ib1
USERCTL=no
ONBOOT=yes
MASTER=bondib0
SLAVE=yes
HOTPLUG=no
BOOTPROTO=none
MTU=65520
```

5. Restart the system.
6. Log in as the `root` user after the system restarts to verify that NIC bonding is running correctly.

```
# cat /proc/net/bonding/bondib0
Ethernet Channel Bonding Driver: v3.2.3 (December 6, 2007)
Bonding Mode: fault-tolerance (active-backup) (fail_over_mac)
Primary Slave: None
Currently Active Slave: ib0
MII Status: up
MII Polling Interval (ms): 100
Up Delay (ms): 5000
Down Delay (ms): 5000
```

```
Slave Interface: ib0
MII Status: up
Link Failure Count: 1
Permanent HW addr: 80:00:00:48:fe:80
```

```
Slave Interface: ib1
MII Status: up
Link Failure Count: 1
Permanent HW addr: 80:00:00:49:fe:80
```

- **Update OpenFabrics Enterprise Distribution on the media server.**

You must use an OpenFabrics Enterprise Distribution (OFED) version that is compatible with the version found in Exadata Database Machine in the media server. You can download the OFED from My Oracle Support [Note 888828.1](#).

- **Configure InfiniBand IPoIB connected mode for best performance.**

No changes are required to the database servers of Exadata Database Machine running Exadata 11g Release 2 (11.2.0.2) and later. However, for custom configurations, you must evaluate the following settings.

The following commands assume a Linux operating system.

1. Verify that Connected Mode is enabled on the system, as follows:

```
# cat /sys/class/net/ib0/mode
connected
# cat /sys/class/net/ib1/mode
connected
```

If the status is “Datagram,” then proceed to step 2 and step 3.

2. Edit the `/etc/ofed/openib.conf` file and search for `SET_IPOIB_CM` and change its value to specify “yes”:

```
# Enable IPoIB Connected Mode
SET_IPOIB_CM=yes
```

3. Reboot the server and re-verify the connected mode again, following the instructions in step 1.

- **Configure MTU Size=65520 on InfiniBand for faster data transmission.**

No changes are required to the database servers of Exadata Database Machine running Exadata 11g Release 2 (11.2.0.2) and later releases. However, for custom configurations, you must evaluate the following settings:

1. Edit the `/etc/sysconfig/network-scripts/ifcfg-ib*` and the `/etc/sysconfig/network-scripts/ifcfg-bondib0` files to add an entry for `MTU=65520`. For example:

```
MTU=65520
```

2. Verify that the MTU size is 65520, as follows:

```
# ifconfig ib0 | grep MTU
UP BROADCAST RUNNING SLAVE MULTICAST MTU:65520 Metric:1
# ifconfig ib1 | grep MTU
UP BROADCAST RUNNING SLAVE MULTICAST MTU:65520 Metric:1
# ifconfig bondib0 | grep MTU
UP BROADCAST RUNNING MASTER MULTICAST MTU:65520 Metric:1
```

3. Reboot the server and verify the MTU size again, following the instructions in step 1.

- **Configure the media server to use the InfiniBand network.**

To direct the backup and restore traffic over the InfiniBand fabric, configure the media management software to favor InfiniBand. Note that each media management software type has its own method of enabling this configuration.

For instance, Oracle Secure Backup has the concept of a **preferred network interface**, which can be set on the media server for a specific list of clients. Other media management software requires this configuration to be defined when the software is installed. See your media management software for information about how to direct traffic over a particular network.

Configuring the Gigabit Ethernet (GigE or 10GigE) Network to Media Server

When connecting the media servers to Exadata Database Machine through Ethernet, connect the eth3 interfaces for GigE or eth4 interface for 10GigE from each database server directly into the data center network. For high availability, multiple network interfaces on the database servers and multiple network interfaces on the media server can be bonded together. In this configuration, configure the eth3 interface as the preferred or primary interface and configure eth2 as the redundant interface for GigE interfaces. For 10GigE configure the eth4 interface as the preferred or primary interface and configure eth5 as the redundant interface.

If throughput is a concern, connect both eth2 and eth3 GigE interfaces from each database server directly into the data center's redundant network. The two interfaces can then be bonded together in a redundant and aggregated way to provide increased throughput and redundancy. The same is possible for 10GigE using the eth4 and eth5 interfaces from each database server.

Follow these best practices:

- **Configure the Gigabit or 10Gigabit Ethernet switch configuration.**

For optimal throughput and availability, configure hardware Link Aggregation in the gigabit switch. The Link Aggregation Control Protocol (LACP)⁹ is defined as part of IEEE 802.1AX-2008 standard. Other software enabled bonding options are available within the operating system of the database servers and media server, which may also be used.

⁹ See http://en.wikipedia.org/wiki/Link_aggregation.

If you are using LACP, then ensure that LACP is supported and configured on the Ethernet switch for Src XOR Dst TCP/UDP Port. See your vendor's Gigabit switch documentation for information about configuring source and destination port load balancing.

On a Cisco 4948 Gigabit Ethernet switch, use the following commands to implement Src XOR Dst TCP/UDP Port:

```
swi-2(config)#port-channel load-balance src-dst-port
swi-2#wr mem
swi-2#sh etherchannel load-balance
EtherChannel Load-Balancing Operational State (src-dst-port):
Non-IP: Source XOR Destination MAC address
  IPv4: Source XOR Destination TCP/UDP (layer-4) port number
  IPv6: Source XOR Destination IP address
swi-2#
```

Additionally, if you are using LACP, then when you configure the `ifcfg-bondeth1` file, change the `BONDING_OPTS` setting to `mode=4`.

- **Configure the database server Gigabit Ethernet.**

No specific changes need to be made to the database servers. However, to obtain higher backup rates, create either a Dual-Port Gigabit Ethernet Configuration or a Dual-Port 10Gigabit Ethernet configuration. See the *Oracle Exadata Storage Server Software User's Guide* for information about bonding multiple interfaces together on Database Server Nodes (database nodes) in Exadata Database Machine.

If you are using LACP, when configuring the `ifcfg-bondeth1` file, change the `BONDING_OPTS` parameter setting to `mode=4`.

- **Configure the media server Gigabit or 10Gigabit Ethernet.**

The following recommendations are applicable only for media servers running Oracle Enterprise Linux Version 5.3 (or later) or RedHat Enterprise Linux Version 5.3 (or later). If your specific media server is running a different operating system, contact your vendor for the appropriate Gigabit configuration.

As with the database server Gigabit Ethernet configuration, no specific changes must be made to the media servers. However, to obtain higher backup rates, create a Multiple-Ported Gigabit or 10Gigabit Ethernet Configuration. The steps to configure bonding on the media server are the same as on the Database Servers. See the *Oracle Exadata Storage Server Software User's Guide* for a detailed procedure.

Configuring Persistent Bindings for Tape Devices

In SAN environments, you must configure persistent bindings so the device address does not change. If the device address changes, the media servers cannot access the device unless you

update the device configuration within Oracle Secure Backup. Therefore, it is very important that your environment maintains consistent device addresses.

Persistent bindings are not configured within Oracle Secure Backup but they are a part of your infrastructure setup. You may configure persistent bindings through the HBA or the operating system. The configuration steps may vary by platform and vendor. See My Oracle Support [Note: 971386.1](#) for an example of creating persistent bindings for device attachments.

Backing up the Oracle Secure Backup Catalog

The Oracle Secure Backup catalog maintains backup metadata, scheduling, and configuration details for the backup domain. Just as it is important to protect the RMAN catalog or control file, the Oracle Secure Backup catalog should be backed up on a regular basis. In Oracle Secure Backup, the catalog backup has been pre-configured:

- **Media family:** OSB_Catalog_MF writes all catalog backups to same tape or tapes.
- **Job summary:** OSB-CATALOG-SUM sends e-mail showing a daily report status of catalog backup to users.
- **Dataset:** OSB-CATALOG-DS defines all directories and files to backup for file system backups.
- **Schedule:** OSB-CATALOG-SCHED shows the schedule for the catalog backup.

The primary catalog backup configuration settings have been defined with only one step remaining which requires user intervention: Edit the OSB-CATALOG-SCHED triggers to specify when to perform the backup.

Disk-Based Backup Strategy

Some of the key benefits to a disk-based backup strategy include:

- Better recovery times for data and logical corruptions and for some Tablespace Point in Time (TSPITR) scenarios
- Ability to using backups directly with no restore

For disk-based backup solutions, Oracle recommends the following:

- Use a fast recovery area
- Perform an initial RMAN level 0 (full) backups
- Perform daily RMAN incremental level 1 backups
- Roll incremental backups into full backup and delay by 24 hours

To scale backup rates for disk:

1. Use all instances and start with two RMAN channels per instance
2. Continue to add another two more RMAN channels for performance.

Optimal backup rates were achieved with all database instances and 2 to 8 RMAN channels per instance. Less than 5% of CPU was utilized for backup operations.

Figure 5 shows the recommended Exadata Storage Server Grid disk group layout. In the figure, the faster (outer) 40% of the disk is assigned to the DATA area, and the slower (inner) 60% of the disk is assigned to the fast recovery area (RECO) area. This can be configured automatically during deployment.

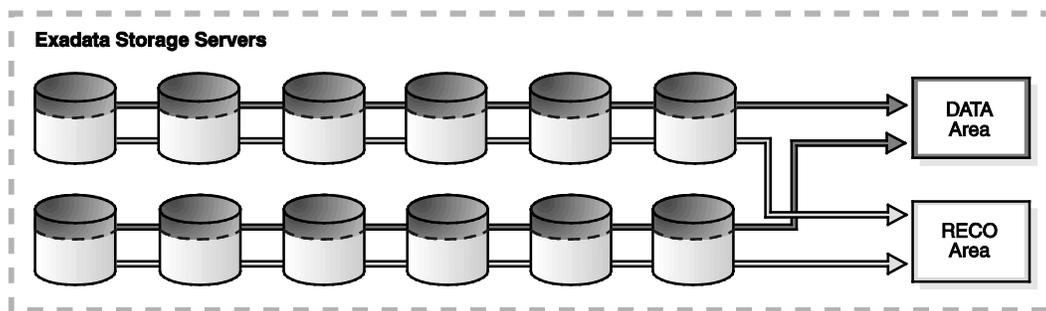
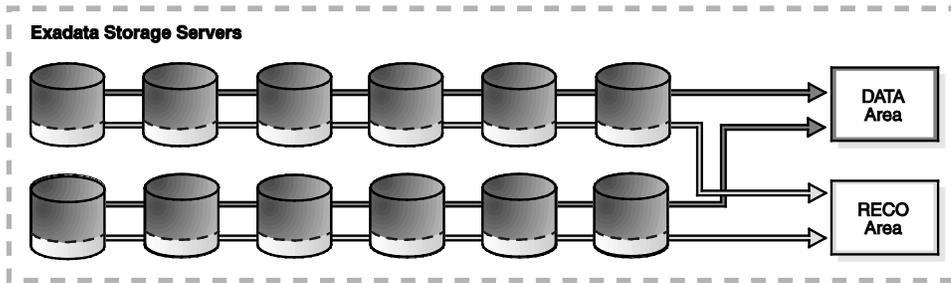


Figure 5: Exadata Storage Server Grid Disk Group Layout for Disk-Based Backup and Recovery

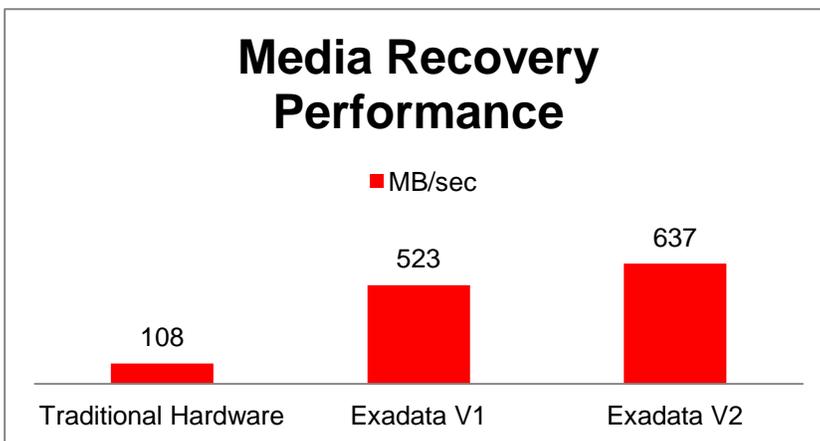
Another strategy is to purchase additional SATA Exadata storage specifically to store the fast recovery area. This allows the application to leverage the full Exadata Database Machine storage grid, allows the use of lower-cost storage for backups, and provides better failure isolation by using separate backup hardware. To reserve more space and bandwidth for the DATA disk group, Oracle recommends using a tape-based backup solution, or at the very least, a hybrid approach where full database backups go to tape and incremental disk backups go to the fast recovery area.

Furthermore for any disk-based backup solution, MAA best practices recommend configuring a high redundancy DATA disk group containing your OCR, Voting, spfiles, data files, redo log groups and control files. Please refer to the following MAA best practices technical whitepaper for ASM configuration best practices on Exadata Database Machine:

<http://www.oracle.com/technetwork/database/features/availability/exadata-maa-131903.pdf>.



When applying redo data on Exadata Database Machine, MAA testing achieved rates greater than 600 MB/sec (or 495 MB/sec with flashback enabled) while applying changes from ETL loads, and greater than 200 MB/sec for OLTP-type workloads. For example, Oracle’s internal production deployment of Beehive Collaborative Application serving over 115,000 Oracle employees applied redo records at 300 MB/sec using Exadata.



Best Practices for Disk-Based Configurations

This section provides:

- [Database Configuration Best Practices](#)
- [RMAN Commands and Configuration](#)

Database Configuration Best Practices

- Use **RMAN incremental backups and block change tracking.**

- To reduce backup time and resources, perform nightly incremental backups to the fast recovery area and merge them into the image copy backup on regular basis. If recovery is needed, then the copies can be directly used as normal data files and recovered to a consistent point, without the need for a restore operation, thus significantly reducing overall recovery time.
- Enable RMAN block change tracking for fast incremental backups. Block change tracking allows RMAN to avoid scanning blocks that have not changed when performing incremental backups. Exadata Storage Server also offloads block inspection from the database servers.

Block change tracking provides the greatest benefit for databases where fewer than 20% of the blocks are changed daily. You may still benefit by using block change tracking with change rates greater than 20%, but testing is recommended to ensure that backup times are reduced.

- **Set the initialization parameter `_file_size_increase_increment=2143289344`**

Set this parameter to optimize the space used when incremental (level 1) backups are taken on the fast recovery area.

- **Reset the initialization parameters `_backup_ksfq_bufsz` and `_backup_ksfq_bufcnt`**

You should reset or remove the `_backup_ksfq_bufsz` and `_backup_ksfq_bufcnt` parameters on systems running Oracle Database release 11.2.0.2 or later releases.

Prior to release 11.2.0.2, it was necessary to explicitly set these parameters to get an optimal backup and restore rate for databases running on Exadata Database Machine. Beginning with Oracle Database release 11.2.0.2, you do not need to explicitly set these parameters because the optimal values are automatically identified for your configuration.

- **Set RMAN configuration setting `FILESERSET=1` when performing incremental backups**

Specify `BACKUP . . . FILESERSET` to specify the maximum number of files in each backup set. A setting of 1 will allow for faster single file database restore operation.

- **Use a backup Oracle Net service for better performance and high availability.**

Configure an Oracle Service to run against all database servers in the cluster. The service is used by the RMAN `BACKUP` command. RMAN automatically spreads the backup load evenly among target the instances offering the service. For example:

```
$ srvctl add service -d <db_unique_name> \
-s <service_name> \
-r <list of preferred instances>
$ srvctl start service -d <db_unique_name> \
-s <service_name>
```

```
Connect to RMAN using the service name:  
$ rman target sys/<sys_password>@<service_name>
```

Note that in order to get true load balancing, refer to Appendix-A

- **Use two to eight RMAN channels per instance.**

In most cases, two RMAN channels per database server are sufficient. Exadata Database Machine X2-8, however, may require eight RMAN channels per instance for the highest throughput. During backup operations, sufficient CPU resources are available for production usage because less than 5% CPU utilization is required for backups. Listener load balancing distributes the connections between the two instances.

- **Set `DB_RECOVERY_FILE_DEST_SIZE` to bound space in the fast recovery area.**

The database writes archived redo log files and any additional recovery files to the fast recovery area. These include any disk backup files such as level 0 image copies and level 1 backup sets as well as Flashback log files (if Flashback Database is enabled). It is important that you set the value of this parameter to less than the total free space in the disk group, which must take account of at least one disk failure and preferably one Exadata Cell failure.

Additionally, if multiple databases are sharing the fast recovery area, ensure that the sum of the space allocated to the different databases is less than the free space in the disk group.

- **Use an external RMAN recovery catalog repository.**

See the [Oracle Database Backup and Recovery User's Guide](#) for more information about the RMAN repository.

RMAN Configuration Commands and Backup Script

- Use the following configuration commands to parallelize backups across all database nodes allowing all the disks and network connections to be leveraged for increased performance. Use two to eight RMAN channels per instance.

RMAN configuration changes

```
CONFIGURE DEFAULT DEVICE TYPE TO DISK;  
CONFIGURE DEVICE TYPE DISK PARALLELISM 16;
```

- Use the following backup script to automate RMAN backups and automation of applying the previous incremental backups.

RMAN script:

```
run {  
  recover  
  copy of database
```

```

        with tag 'Disk_Backup';
    backup
        incremental level 1
        for recover of copy
        with tag 'Disk_Backup'
        database;
}

```

Restore and Recovery Best Practices

Restore rates of up to 14 TB/hour were achieved when writing to high redundancy DATA disk group. Writing to normal redundancy disk group achieved 17 TB/hour and restoring to existing files in a normal redundancy disk group achieved 24 TB/hour. Create a restore service that runs across all database instances and use two to eight RMAN channels per database instance.

For example,

1. Create an Oracle service for restore for all available database nodes.

The service is used by the RMAN restore command. RMAN automatically balances the restore load among the targeted instances. For example:

```

$ srvctl add service -d <db_unique_name> \
-s <service_name> \
-r <list of preferred instances>
$ srvctl start service -d <db_unique_name> \
-s <service_name>

```

Note that in order to get true load balancing, refer to Appendix-A

2. Connect to RMAN using the service name:

```
$ rman target sys/<sys_password>@<service_name>
```

- o For tape-based channels = 8 tape drives


```

CONFIGURE DEFAULT DEVICE TYPE TO SBT;
CONFIGURE DEVICE TYPE SBT PARALLELISM 8;

```
- o For tape-based channels = 14 tape drives


```

CONFIGURE DEFAULT DEVICE TYPE TO SBT;
CONFIGURE DEVICE TYPE SBT PARALLELISM 14;

```
- o For disk-based channels = n times 2 where $n = <the\ number\ of\ instances>$

```

CONFIGURE DEFAULT DEVICE TYPE TO DISK;
CONFIGURE DEVICE TYPE DISK PARALLELISM 16;

```

3. Issue the following RMAN command:

```
RUN {  
  restore database;  
  recover database;  
}
```

4. To restore the Oracle Secure Backup catalog, see the Oracle Secure Backup documentation at

http://download.oracle.com/docs/cd/E14812_01/doc/doc.103/e12834/catalog_recovery.htm#BABIJEIH

Offload Backups with Oracle Data Guard

Oracle Data Guard is the Oracle MAA prescribed disaster-recovery solution to protect mission critical databases residing on Exadata Database Machine and the Exadata Storage Server. Data Guard physical standby databases support all Oracle datatypes and features, including Exadata Hybrid Columnar Compression, and are able to support the very high transaction volume driven by Exadata Database Machine.

Data Guard standby databases are also used to offload backups from a production database (primary). Both disk and tape-based backups can be performed using a physical standby database. Oracle Active Data Guard (an extension of Data Guard functionality) can be used to offload fast incremental backups (RMAN block change tracking), further reducing backup times and the impact to the primary database. Additional benefits of Active Data Guard include offloading read-only queries and reports from the primary to a synchronized physical standby database and automatic block repair should Oracle detect a block corruption. All Data Guard standby databases can also be used to detect lost write corruptions and for database rolling upgrades, and other maintenance, while also providing disaster protection.

For more information, see the “Oracle Data Guard: Disaster Recovery for the Exadata Database Machine” white paper on the MAA Web site at

<http://www.oracle.com/technetwork/database/features/availability/maa-wp-dr-dbm-130065.pdf>

Caution when Considering non-Exadata Storage for Backups.

MAA best practices provided in this paper have focused on configurations using Exadata Storage Servers as the backup target. Extra care must be taken when considering use of non-Exadata storage solutions for backups. Differences in network connectivity, storage configuration, and protocols used by non-Exadata storage will impact backup performance and complexity. Non-Exadata storage cannot support all Oracle features (for example there is no benefit of storage cell

offload for RMAN block-change tracking). For these reasons, MAA best practices recommend using Exadata storage for optimal backup performance.

However, there may be cases where users may have no choice other than to utilize non-Exadata storage for disk-based backups. In such cases, then:

- Perform database backups to a NAS appliance, such as Sun ZFS Storage Appliance, that is connected to Exadata Database Machine via IP-based protocols using 1GigE, 10GigE or InfiniBand interfaces. The Sun ZFS Storage Appliance architecture is optimized for backup operations using NFS. For technical details on configuration and use of the Oracle Sun ZFS Storage Appliance see <http://www.oracle.com/technetwork/server-storage/sun-unified-storage/overview/unified-storage-overview-165517.html>. See also (link to the ZFS backup & recovery paper)
- For backing up to storage not directly connected to Exadata Database Machine (e.g. SAN storage), it is necessary to use an intermediate server that acts as an iSCSI or NFS server since a direct SAN or Fibre Channel connection over Ethernet is not supported. Note that most standard operating system NFS servers have poor performance for synchronous writes making them unsuitable for applications on Exadata Database Machine when fast backup performance is required. When using an NFS server it must be understood that backup performance will be significantly less than what the storage array should be able to achieve, and significantly less than can be achieved by direct backup to Exadata Storage Server.
- Thorough performance testing is recommended to insure that backup, restore and recovery service level agreements are met.
- Generic tuning and performance troubleshooting recommendations are provided in My Oracle Support Note 1275894.1

Monitoring and Troubleshooting

Monitoring RMAN

When the RMAN job is executed the job transcript is written to `stdout` by default, but the output can be redirected to a log file that can be analyzed for errors and warnings as well as to review backup piece names that are written to. Additionally, RMAN uses the `NLS_DATE_FORMAT` environment variable to report times in hours / minutes and seconds, that can be useful to monitor run times.

Monitoring and Troubleshooting Oracle Secure Backup

Where you begin troubleshooting depends on the problem reported:

- Primary Oracle Secure Backup resources:
 - Backup/restore Job transcript and/or properties
 - Daemon (process) logs
 - Device logs
- External environmental areas to review:
 - Operating System configuration settings
 - Confirm Oracle Secure Backup user has the correct Operating System privileges to perform backup and restore operations
 - Confirm the tape device is accessible to the host

Job Transcripts

Job transcripts are usually the first place to begin providing detailed error messages for Oracle Secure Backup jobs. Job transcripts may be viewed using the Web Tool or command-line interface, `obtool`.

Obtool Commands:

```
ob> lsjob -A           Listing of all jobs useful to obtain Job ID
ob> catxcr -f10 <jobid>  Displays entire transcript for the job
```

Job Properties

Job properties are another resource to determine what caused corresponding job issue. In some circumstances the job transcripts may not contain information such as when:

- Backup/restore job failed before it began
- Parent job scheduled when the backup/restore request issued but child job, which actually transferred data never, began

Job Properties and Logs

Job properties and logs may provide information describing why the job failed and are available from the Web Tool or the command line. To access properties from the command line, use the command:

```
ob> lsjob --log <jobid>
```

Monitoring TCP/IP Traffic

Oracle Secure Backup sends the data across the TCP/IP stack. To verify the backup rates, you can view the “`sar`” output either in real time or historically to see the transfer rates achieved from the database servers to the media servers or vice versa.

Conclusion

Exadata Database Machine used with MAA best practices described in this white paper allow you to backup, restore, and recover Oracle Database extremely fast, using either disk-based backup or tape-based backup and restore mechanisms. Exadata Database Machine and Exadata Storage Server benefit from their native integration with Oracle Database and other Oracle high availability features. MAA best practices provide an optimal solution for Oracle databases of any size and recommendations are straightforward to implement using standard RMAN commands.

Appendix A – Achieving 100% load balancing for RMAN backup scripts

In order to achieve 100% load balancing, services must be created that default to the preferred node, but for high availability may fail over to alternate nodes on the cluster where the database is running. Additionally, the RMAN channels must be pointed to by the individual services.

- 1) Create a service that runs on a specific nodes in the cluster:

```
srvctl add service -d <dbname> -s <service name #1> -r <instance1>
-a <instance2>,<instancen>
srvctl add service -d <dbname> -s <service name #2> -r <instance2>
-a <instance1>,<instancen>
```

For example

```
srvctl add service -d dbm -s bkup1 -r dbm1 -a
dbm2,dbm3,dbm4,dbm5,dbm6,dbm7,dbm8
srvctl add service -d dbm -s bkup2 -r dbm2 -a
dbm1,dbm3,dbm4,dbm5,dbm6,dbm7,dbm8
srvctl add service -d dbm -s bkup3 -r dbm3 -a
dbm1,dbm2,dbm4,dbm5,dbm6,dbm7,dbm8
srvctl add service -d dbm -s bkup4 -r dbm4 -a
dbm1,dbm2,dbm3,dbm5,dbm6,dbm7,dbm8
srvctl add service -d dbm -s bkup5 -r dbm5 -a
dbm1,dbm2,dbm3,dbm4,dbm6,dbm7,dbm8
srvctl add service -d dbm -s bkup6 -r dbm6 -a
dbm1,dbm2,dbm3,dbm4,dbm5,dbm7,dbm8
srvctl add service -d dbm -s bkup7 -r dbm7 -a
dbm1,dbm2,dbm3,dbm4,dbm5,dbm6,dbm8
srvctl add service -d dbm -s bkup8 -r dbm8 -a
dbm1,dbm2,dbm3,dbm4,dbm5,dbm6,dbm7
```

- 2) Start the services by running

```
srvctl start service -d <dbname> -s <service_name #1>
```

Alternatively, to start all the services for a particular database

```
srvctl start service -d <dbname>
```

- 3) When running RMAN, use the service name in the connect string for the individual channels allocated. The RMAN script will look similar to

```
run {
allocate channel ch01 connect string
'sys/<password>@<scan_address>/<service_name>;
allocate channel ch02 connect string
'sys/<password>@<scan_address>/<service_name>;
...
```

```
allocate channel ch16 connect string
'sys/<password>@<scan_address>/<service_name>;
backup database ...;
}
```

For example

```
run
{
allocate channel ch01 device type disk connect 'sys/welcome1@dbm-
scan/bkup1';
allocate channel ch02 device type disk connect 'sys/welcome1@dbm-
scan/bkup2';
allocate channel ch03 device type disk connect 'sys/welcome1@dbm-
scan/bkup3';
allocate channel ch04 device type disk connect 'sys/welcome1@dbm-
scan/bkup4';
allocate channel ch05 device type disk connect 'sys/welcome1@dbm-
scan/bkup5';
allocate channel ch06 device type disk connect 'sys/welcome1@dbm-
scan/bkup6';
allocate channel ch07 device type disk connect 'sys/welcome1@dbm-
scan/bkup7';
allocate channel ch08 device type disk connect 'sys/welcome1@dbm-
scan/bkup8';
allocate channel ch09 device type disk connect 'sys/welcome1@dbm-
scan/bkup1';
allocate channel ch10 device type disk connect 'sys/welcome1@dbm-
scan/bkup2';
allocate channel ch09 device type disk connect 'sys/welcome1@dbm-
scan/bkup1';
allocate channel ch10 device type disk connect 'sys/welcome1@dbm-
scan/bkup2';
allocate channel ch11 device type disk connect 'sys/welcome1@dbm-
scan/bkup3';
allocate channel ch12 device type disk connect 'sys/welcome1@dbm-
scan/bkup4';
allocate channel ch13 device type disk connect 'sys/welcome1@dbm-
scan/bkup5';
allocate channel ch14 device type disk connect 'sys/welcome1@dbm-
scan/bkup6';
allocate channel ch15 device type disk connect 'sys/welcome1@dbm-
scan/bkup7';
allocate channel ch16 device type disk connect 'sys/welcome1@dbm-
scan/bkup8';
backup
    as backupset
    incremental level 0
    section size 128g
    database;
}
```

Appendix B – Test Environment

The MAA test environment used in this white paper was comprised of the following hardware:

- One Exadata Database Machine
- One Sun Fire X4170 server for the Oracle Secure Backup administrative server
- Two Sun Fire X4275 servers for Oracle Secure Backup media servers
- One Sun StorageTek SL3000 with 8 Sun StorageTek T10000C tape drives and 8 LTO4 tape drives
- One Brocade 5100 8Gbit Fiber Channel switch



Backup and Recovery Performance and Best
Practices for Exadata Cell and the Oracle
Exadata Database Machine

February 2013

Author: Andrew Babb

Contributing Authors: Lawrence To, Steve Fried,
Steve Wertheimer, Douglas Utzig, Michael
Nowak, Tim Chien, Viv Schupmann

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



| Oracle is committed to developing practices and products that help protect the environment

Copyright © 2011, Oracle and/or its affiliates. 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.