ORACLE
UNIVERSITY

Integrated Cloud Applications & Platform Services

Exadata Database Machine: 12*c* Administration Workshop

Student Guide - Volume II

D92887GC20

Edition 2.0 | November 2016 | D97406

Learn more from Oracle University at **education.oracle.com**

ORACLE

Author

Peter Fusek

Technical Contributors and Reviewers

Akshay Shah
Alex Tsukerman
Amit Ganesh
Andrew Babb
Aslam Edah-Tally
Barb Lundhild
Bharat Baddepudi
Bill Hodak
Boris Erlikhman
Branislav Valny
Bruce Kyro
Bruno Dagostini
Bundit Jianpinitnan
Caroline Johnston
Christian Craft
Dan Norris
Darryl Presley
Dave Winter
David Hernandez Mendoza
David Hitchcock
Deba Chatterjee
Douglas Utzig
Ed Gilowski
Eric Siglin
Francois Marcoux
Georg Schmidt
Gerlinde Frenzen
Harald van Breederode
Ira Singer
James He
James Womack
Jean-Francois Verrier
Jerry Lee
Jia Shi
Jignesh Patel
Jim Hall
Jim Spiller
Jim Viscusi
Joe Fu
Joel Goodman
Juan Loaiza
Kam Shergill
Kevin Jernigan
Kodi Umamageswaran
Krishnanjani Chitta
Lachlan Williams
Larry Justice
Lawrence To
Louis Nagode
Mahesh Subramaniam
Maria Billings
Maria Colgan
Mark Fuller
Mark Scardina
Mark Van de Wiel
Marshall Presser
Martin Jensen
Michael Cebulla
Michael Nowak
Naoki Kato
Nikolaus Dilger
Nilesh Choudhury
Ravindra Dani
Raymond Dutcher
Richard Exley
Robert Carlin
Robert Pastijn
Roger Hansen
Sabyasachi Banerjee
Sean Kim
Selcuk Aya
Scott Gossett
Sriram Palapudi
Steven Lemme
Sue Lee
Sugam Pandey
Sumeet Lahorani
Sundararaman Sridharan
Suresh Gogune
Tim Shetler
Trevor Bowen
Umesh Panchaksharaiah
Uwe Hesse
Varun Malhotra
Vern Wagman
Vijay Sridharan
Vikram Kapoor
Vimala Jacob

Editors

Arijit Ghosh
Aishwarya Menon

Graphic Designer

Seema Bopaiah

Publishers

Sujatha Nagendra
Sumesh Koshy
Asief Baig

Contents

10

Consolidation Options and Recommendations

ORACLE

Objectives

After completing this lesson, you should be able to:

- Describe the options for consolidating multiple databases on Exadata
- Explain the benefits and costs associated with the different options

ORACLE

Consolidation: Overview

- Why consolidate:
  - To minimize idle resources
  - To maximize efficiency
  - To lower costs
- Consolidation examples:
  - Hosting multiple virtual servers on one physical server
  - Hosting multiple database instances on one server
  - Supporting different users on one database instance

ORACLE

Consolidation is one of the major strategies that organizations of all types are pursuing to achieve greater efficiencies in their operations. In essence, consolidation allows organizations to increase the utilization of IT resources so that idle cycles can be minimized. This, in turn, lowers costs because fewer resources are required to achieve the same outcome. For example, systems that experience peak load at different times of the day can share the same hardware, rather than use dedicated hardware, which might be idle during non-peak periods.

Consolidation can be achieved in different ways depending on the systems and circumstances involved. Running multiple virtual machines and multiple applications on a single server are both examples of consolidation. Using one Oracle Database instance to support multiple, different user populations, potentially from different organizations or lines of business, is another example of consolidation.

The performance and capacity of Exadata makes it an ideal consolidation platform. This lesson outlines typical approaches for consolidation on Exadata, along with associated recommendations, benefits and costs.

Different Consolidation Types

- Database:
  - Multiple databases consolidated on to a single platform
  - Separate instances for each database
  - Separate database management
  - Platform resources are shared by database instances
- Application Schema:
  - Multiple application schemas consolidated into a single database
  - One database instance (or one set of Real Application Clusters [RAC] instances)
  - One database to manage

ORACLE

The slide outlines two types of consolidation typically associated with Oracle Database.

Database consolidation is the practice of hosting multiple databases on a single platform, such as Exadata. With database consolidation, a degree of autonomy is maintained for each database. Each database uses separate instances and each database is managed separately. However, all the databases must share the resources of the underlying platform.

Application schema consolidation is the practice of hosting multiple application schemas in a single database. The main benefit of this style of consolidation is that the overhead of multiple database instances can be avoided. The main drawback is that the application schemas become inexorably linked, which can cause difficulties in certain circumstances. For example, if one application could benefit from a database upgrade and the other application is not certified against the upgraded database version, then an inevitable compromise is required. Also, there are circumstances where application schemas can clash, making consolidation impractical. For example, if two applications contain public synonym definitions with the same name, the clash could compromise one of the applications.

Different Consolidation Types

- Virtual Machine:
  - Multiple virtual machines consolidated on to a single platform
  - One or more databases hosted on each virtual machine
  - Platform resources are shared by multiple operating system and database instances
  - Separate operating system and database management
- Oracle Multitenant Architecture:
  - New database architecture in Oracle Database 12*c*
  - Multiple pluggable databases consolidated into a single container database
  - One container database and instance set to manage
  - Complete logical separation between pluggable databases

ORACLE

The slide outlines another two consolidation types typically associated with Oracle Database.

Consolidation using Virtual Machines (VMs) is the practice of running multiple VMs on a single server. VMs can improve server utilization while maintaining separation between databases and their underlying operating system environments. However, the effectiveness of this approach is limited because of the resource footprint consumed by multiple operating system and database instances. Also, this approach does nothing to reduce the number of operating system and database instances that need to be managed.

Oracle Multitenant is a new architecture introduced in Oracle Database 12*c*. It introduces the concept of a container database (CDB), which can support multiple pluggable databases (PDBs). Consolidation is achieved by plugging-in multiple PDBs into a single CDB. A CDB has a single set of background processes and shared memory area (SGA) that is shared by all of the PDBs. However, the PDBs are logically isolated from each other just like independent databases.

The remainder of the lesson discusses each of the consolidation types in more detail.

Core Principles for Database Consolidation

- Consolidate databases with similar availability and planned maintenance objectives.
- Other considerations include performance, system requirements, security, and organization boundaries.
- Consolidating applications onto Exadata is possible but not recommended.

Exadata is well suited as a consolidation platform for Oracle databases. The grid architecture, which supports Exadata storage and the Oracle Real Application Clusters software, provides many potential configuration options for consolidation.

When considering database consolidation on Exadata, a few core principles should be applied. Firstly, the availability and planned maintenance objects of the databases that are candidates for consolidation should be similar. If this is not the case, inevitably one of the systems will be compromised in some way. For example, if a critical business system shared the same environment as a development system, either business operations would be disrupted if the system is patched and rebooted, or development activities might be interrupted if the patch is not applied.

Other considerations, such as performance objectives, system requirements, security, and organization boundaries must also be considered to ensure compatibility between the various consolidation candidates. For example, if a security directive mandates physical separation for data in two different systems, then obviously they would not be good candidates for consolidation.

Finally, although it is technically feasible to host non-database applications onto the Exadata database servers, this is not recommended mainly because altering the Exadata software to meet the application requirements could degrade the performance and availability of Exadata.

Recommended Consolidation Approach

- Categorize databases into different groups based on availability and recovery objectives:
  - **Platinum:** Maximum availability with a goal of no outages for platinum-ready applications
  - **Gold:** Comprehensive high availability and disaster recovery
  - **Silver:** High availability with automatic failover
  - **Bronze:** Single-instance database
- Create hardware pools that consolidate databases within each group but not across groups:
  - Keep each hardware pool a manageable size

ORACLE

The following approach is recommended when performing database consolidation by using Exadata. First, categorize the databases into different groups based on similar objectives as discussed on the previous page. The groupings that are listed in the slide are based on the Oracle Maximum Availability Architecture (MAA) high availability reference architectures. Each reference architecture encompasses different technologies and recommendations to deliver the appropriate service level. For more information see www.oracle.com/technetwork/database/availability/maa-reference-architectures-2244929.pdf.

Then in each group, consolidate the databases on the same Exadata environment. To ensure consistent terminology throughout this lesson, each separate consolidated environment is referred to as a hardware pool. It is not recommended to consolidate databases from more than one group on the same hardware pool.

It is also recommended to limit the size of a hardware pool to a small number of Exadata racks. This should provide ample space and power for most requirements. Where greater capacity is required, multiple hardware pools should be considered because the complexity of managing an extremely large environment starts to offset the consolidation benefits.

Individual databases requiring more resources than a hardware pool are probably not a good candidate for consolidation and should be maintained in their own environment.

Recommended Storage Configuration for Consolidation

- One shared Exadata storage grid per hardware pool:
  - Use the default disk groups:
    - DATA for database files, RECO for FRA, DBFS_DG for clusterware shared files and DBFS
    - Stripe disk groups across all cells and all disks.
  - Use high redundancy for DATA.
  - Use IORM to manage and allocate I/O resources.
  - Set the `COMPATIBLE.RDBMS` disk group attribute to the minimum database software version being used.
- Benefits:
  - Standard configuration: Is the easiest to manage
  - Balanced configuration: Provides high availability and performance, and suits most circumstances

The recommended storage configuration for consolidation is to use one shared ASM storage grid for each consolidated environment (hardware pool). The default disk group configuration resulting from the installation and configuration process should be used. By default, the DBFS_DG disk group is used to house the clusterware shared files (cluster registry and voting files) and it can also be used to house a file system by using the Database File System (DBFS) feature. Use the DATA disk group to store database files and the RECO disk group to mainly store the fast recovery area (FRA). During initial configuration, the DATA and RECO disk groups are configured to use equal-sized disk stripes (grid disks) on every disk in every Exadata cell.

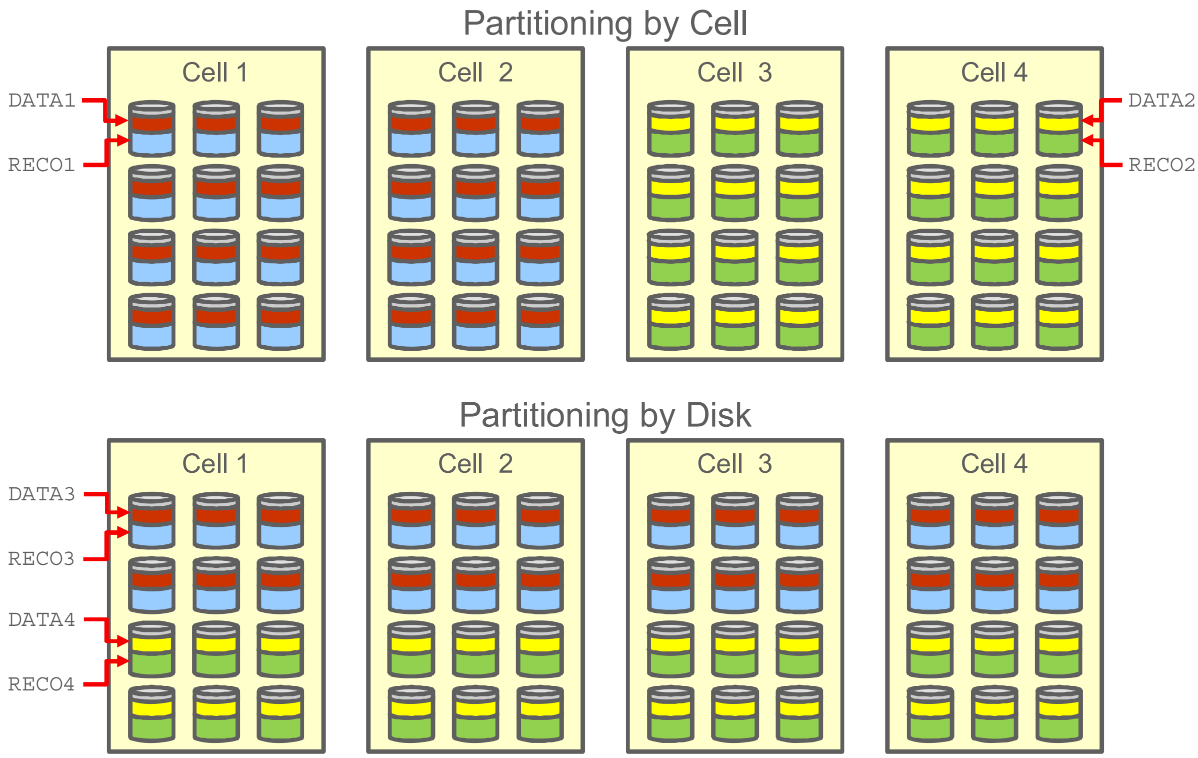
Where possible, it is recommended to use the high redundancy setting for the DATA disk group. This provides maximum protection from cell outages and storage failures.

For consolidated environments, Exadata I/O Resource Management (IORM) is a key tool. It provides a flexible way of managing I/O resource to ensure that database workloads are prioritized in line with specified policies. Note that at this time, IORM does not prioritize writes; therefore, special care must be taken when consolidating write-intensive applications.

The recommended storage configuration has the following major benefits:

- It is conceptually simple and inherently easy to manage. It is also the most familiar and supportable configuration.
- It is a balanced configuration, which is suitable for most circumstances and provides high levels of availability and performance.

Alternative Storage Configurations



The diagrams in the slide illustrate alternative storage configuration strategies.

The diagram at the top of the slide illustrates cell-level storage partitioning. With this approach, different sets of disk groups, are configured on separate cell groups. In the illustration, one set of disk groups (`DATA1` and `RECO1`) is stored on the first two cells, and another set of disk groups (`DATA2` and `RECO2`) is stored on a separate set of cells. This strategy is employed when you use the Oracle Exadata Deployment Assistant to define multiple physical clusters on a single Exadata rack. The basic idea of cell-level partitioning can be extended to accommodate more disk group sets, or varied so that the DATA and RECO disk groups occupy separate cells.

The diagram at the bottom of the slide illustrates partitioning Exadata storage by disks within cells. With this approach, different disk groups are configured on a subset of disks within each cell. In the illustration, one set of disk groups (`DATA3` and `RECO3`) is stored on half of the disks in each cell, and another set of disk groups (`DATA4` and `RECO4`) is stored on the remaining disks in each cell. This idea could be extended to accommodate more disk group sets, or varied so that different disk groups occupy a specific number of disks on each cell.

Both approaches can also be implemented in conjunction with database-scoped Exadata storage security to provide robust storage separation for different databases. Exadata storage security is discussed in the lesson titled *Exadata Storage Server Configuration*.

You can also create other alternative storage configurations by combining the partitioning approaches outlined in this slide. However, note that the only way to implement partitioning by disk, or any combined approach, is by manually reconfiguring Exadata storage, which would result in a highly non-standard configuration and is generally not recommended.

Benefits and Limitations of Partitioned Storage Configurations

- Benefits:
  - ✓ Storage isolation
    - Different databases occupy separate cells or disks.
  - ✓ Patch isolation
    - Different sets of cells can be patched to different levels.
- Limitations:
  - × Manual configuration
    - Partitioned configurations need to be hand-crafted.
  - × Reduced I/O bandwidth
    - A single operation cannot leverage the entire storage grid.
  - × Additional management overhead
    - Commands need to be duplicated.
    - Special adjustments may be required.

ORACLE

The alternative storage configuration options have the following benefits and limitations:

- ✓ Partitioning disk groups across physically separate cells or disks provides better isolation between the activities on the different disk groups. For example, this can be beneficial in cases where security policies mandate physical separation of data, or where customers desire to isolate the I/O from different databases.
- ✓ Using the approach of partitioning by cells, different cells can be patched to different levels. This patch isolation can be beneficial in circumstances where a patch is required for one database but has undesirable consequences for another database. Note that this approach is limited to a degree because of the dependencies that exist between different Exadata components.
- × Non-standard storage configurations must be manually configured, which requires additional planning and effort compared with the recommended configuration.
- × Any alternative configuration where disk groups are not striped across all the disks in all the cells must have a reduced I/O bandwidth capability compared with the recommended standard storage configuration.
- × Non-standard storage configurations require additional effort to manage them. At the very least, the existence of more disk groups increases the number of commands required to manage them. In addition, management operations that assume a standard storage configuration need to be specially altered to work in a non-standard environment. For example, the instructions supplied with a patch might need adjustment to suit a non-standard storage configuration.

Cluster Configuration Options

- Recommended configuration:
  - Use one cluster per hardware pool.
  - Use database services to manage database workloads across the available database servers.
  - Use Oracle Data Guard to protect against cluster failure for critical hardware pools.
- Alternative configuration:
  - Configure multiple clusters per hardware pool:
    - ✓ Better isolation for cluster failure and patching
    - × Additional management effort and complexity
    - Requires a separate set of disk groups for each cluster
    - A minimum of three nodes per cluster is recommended.

ORACLE

The recommended cluster configuration for Exadata is to use one cluster for each environment (hardware pool). For consolidation environments, database services are recommended as a primary means of managing database workloads across the available servers. For critical hardware pools, Oracle Data Guard is recommended to protect against cluster failure.

Multiple clusters running on a single Database Machine is supported. In this mode, each cluster occupies a separate subset of database servers. This alternative can provide better isolation if the cluster fails or if the Grid Infrastructure software requires patching; however, this comes at the cost of additional management effort and complexity.

A non-standard cluster configuration also implies a non-standard storage configuration because disk groups cannot be shared across different clusters. Refer to the benefits and limitations associated with non-standard storage configurations before considering a non-standard cluster configuration.

Finally, if you choose to implement a non-standard cluster configuration, it is recommended to have a minimum of three nodes in each cluster.

Operating System Parameter Recommendations

- Set the number of shared memory segments (SHMMNI) greater than the number of databases.
- Set the maximum shared memory segment size (SHMMAX) to 85% of physical memory size.
- Set the maximum total number of system semaphores (SEMMNS) greater than the sum of all database processes.
- Set the maximum number of semaphores in a semaphore set (SEMMSL) greater than the largest number of processes in any single database.
- For Linux, if PageTables in /proc/meminfo is > 2% of physical memory size, consider setting HugePages equal to the sum of all shared memory segments:
  - See My Oracle Support notes 401749.1 and 361468.1.

ORACLE

The following are the recommended Linux operating system parameter settings for consolidation:

- Set the number of shared memory segment identifiers greater than the number of databases running on the node. Check the setting for `kernel.shmmni` by using the `sysctl` command. If necessary, make adjustments by setting `kernel.shmmni` in `/etc/sysctl.conf`.
- Set the maximum shared memory segment size to 85% of the database server physical memory size. Check the setting for `kernel.shmmax` by using the `sysctl` command. If necessary, make adjustments by setting `kernel.shmmax` in `/etc/sysctl.conf`.
- Set the maximum number of semaphores in the system greater than the sum of the processes for all of the databases running on the system. Check the setting for `kernel.sem` by using the `sysctl` command. The setting for `kernel.sem` contains a list of four numeric values. The setting for the total number of semaphores on the system, also known as SEMMNS, is the second value in the list. If necessary, make adjustments by setting `kernel.sem` in `/etc/sysctl.conf`.

- Set the maximum number of semaphores in a semaphore set greater than the largest number of processes for any single database running on the system. Check the setting for `kernel.sem` by using the `sysctl` command. The setting for the maximum number of semaphores in a semaphore set, also known as `SEMMSL`, is the first value in the list. If necessary, make adjustments by setting `kernel.sem` in `/etc/sysctl.conf`.
- HugePages is a Linux kernel feature that enables larger page sizes within the operating system. This, in turn, enables processes referencing large amounts of memory to have smaller page tables, thus reducing page table memory usage. Consider implementing HugePages if `PageTables` in `/proc/meminfo` is > 2% of physical memory. When set, HugePages should at least equal the sum of the shared memory used by all the databases. Because any change to HugePages requires a reboot, you can choose to set HugePages to a value that covers all of the database you plan to implement in the foreseeable future.
  If all the database instances are running, the amount of shared memory being used can be calculated by analyzing the output from the `ipcs -m` command. My Oracle Support note 401749.1 provides a script that can be used to determine the amount of shared memory in use. My Oracle Support note 361468.1 describes how to set HugePages on 64-bit Linux.

Database Memory Recommendations

- Calculate the Oracle Database memory usage as follows:
  - For OLTP: sum((`PGA_AGGREGATE_TARGET` + `SGA_TARGET`) + (4 MB * Maximum Processes))
  - For data warehousing using Oracle Database 11*g*: sum((3 x `PGA_AGGREGATE_TARGET`) + `SGA_TARGET`)
  - For data warehousing using Oracle Database 12*c*: sum(`PGA_AGGREGATE_LIMIT` + `SGA_TARGET`)
- Ensure that the Oracle Database memory usage does not exceed:
  - The database server memory size for non-critical hardware pools
  - 75% of the database server memory size for critical hardware pools
- Monitor memory usage and maintain a free memory target of at least 5% at all times.

ORACLE

The general recommendation is that the instance memory areas (SGA plus PGAs) for all of the databases in a consolidation setting should occupy less than the physical memory on each database server. For critical hardware pools, Oracle recommends a more conservative approach, that is, the instance memory areas should occupy less than 75% of the physical memory on each database server. The formulas in the slide provide general guidance on how to measure the instance memory usage for OLTP and data warehousing applications.

Note that the `PGA_AGGREGATE_TARGET` database parameter setting does not enforce a hard maximum for PGA usage. For some data warehouse applications, three times the specified target has been observed. For OLTP applications, the potential overrun is generally much less. Commencing with Oracle Database 12*c*, the `PGA_AGGREGATE_LIMIT` database parameter provides a mechanism to place a hard limit on PGA memory usage.

Exceeding the recommended memory usage can potentially cause performance problems. If adjustments are required to fit existing databases under the recommended ceiling, then you must ensure that the resulting values are sufficient for the applications using the associated databases. Otherwise, you might need to move some databases to another hardware pool.

Monitoring of application and system memory utilization is required to ensure that there is sufficient memory throughout the workload and business cycles. Oracle recommends a free memory target of at least 5% at all times.

CPU Management Recommendations

- Use instance caging to prevent individual databases from consuming too much CPU.
- Recommended `CPU_COUNT` parameter settings:
  - For all databases: `CPU_COUNT` >= 2
  - For critical pools hardware pools:
    sum(`CPU_COUNT`) <= the total number of CPU cores
  - Otherwise:
    sum(`CPU_COUNT`) <= three times the total number of CPU cores
- Use the database resource manager to control CPU allocation within each database.
  - Use the `MAX_UTILIZATION_LIMIT` directive attribute to limit the CPU utilization for consumer groups.

ORACLE

Instance caging is an important tool for managing CPU resources in consolidation environments where each database server hosts multiple database instances. It uses the `CPU_COUNT` initialization parameter to specify the maximum number of CPU cores that a database can simultaneously use.

Instance caging can be used to guarantee that instances do not compete with each other for CPU resources. It can also be configured so that CPU allocations are over-subscribed.

For example, a server with 32 CPU cores could host 4 databases each having a `CPU_COUNT` setting of 16. This configuration is reasonable in cases where the databases are mostly idle or have different usage patterns. It guarantees that no single database can flood the entire system. However, if all the databases become heavily loaded at the same time, then obviously there would be contention for CPU resources and degraded performance.

The following are the recommended guidelines for setting the `CPU_COUNT` parameter:

- For all databases, it is recommended to set `CPU_COUNT` to at least 2 because setting `CPU_COUNT=1` can make database instances more susceptible to hangs or poor performance. In situations where database instances require fewer than 2 full CPUs, it is acceptable to use an over-subscribed approach.

- For critical hardware pools, the aim is to ensure that databases are never starved of CPU resources. To achieve that aim, the total of the `CPU_COUNT` parameter values across all the databases should be less than 100% of the total number of CPU cores on the server. This policy is particularly beneficial for CPU-intensive databases, such as those executing large analytical queries or large volumes of PL/SQL processing.
- For less critical hardware pools, over-provisioning the CPU resources is a valid approach. A `CPU_COUNT` total of up to three times the available number of CPU cores is reasonable. Over-provisioning is particularly effective for databases that are not CPU-intensive, or databases having different active and idle times, Remember that you should monitor the environment and can dynamically adjust `CPU_COUNT` settings to cater to different processing cycles.

Note that instance caging is enabled by setting the `CPU_COUNT` initialization parameter and enabling the database resource manager by using the `RESOURCE_MANAGER_PLAN` initialization parameter. You can set `RESOURCE_MANAGER_PLAN` to use the predefined plan `DEFAULT_PLAN` as a basic starting value.

Database resource manager is also recommended for managing CPU resources within each database. The database resource manager can be configured to share CPU resources between database workloads. By default, the resource manager only arbitrates to allocate CPU resources when the utilization increases to the point of contention. In other words, a workload can consume all the available resources if no other workload is competing for them. However, if you want to set an upper limit on the CPU share allocated to a particular workload, you can do so by using the `MAX_UTILIZATION_LIMIT` resource manager directive.

Process Management Recommendations

- Adjust PROCESSES for ASM:
  - For 2 to 10 databases: PROCESSES = 50 + (50 * number of databases)
  - For 11 or more databases: PROCESSES = 450 + (10 * number of databases)
- Limit PARALLEL_MAX_SERVERS:
  - For 2-socket servers: sum(PARALLEL_MAX_SERVERS) <= 240
  - For 8-socket servers: sum(PARALLEL_MAX_SERVERS) <= 1280
- Limit Data Guard redo apply parallelism to 16 or less.
- Process and connection guidelines:
  - Limit active processes:
    - For average applications: Active processes <= 5 x CPU_COUNT
    - For very CPU-intensive applications: Active processes <= CPU_COUNT
  - Use connection pools to manage application connections.
  - Use database shared server processes to reduce the process count.
  - Use the connection rate limiter to prevent connection storms.

ORACLE

The following process management recommendations apply for database consolidation:

- Using the formulas in the slide, adjust the PROCESSES parameter for ASM. For example, if you are consolidating five databases, set PROCESSES=300 in your ASM instance parameter file.
- Ensure that PARALLEL_MAX_SERVERS parameter for each database is set so that the sum of all the parameters does not exceed the recommended maximums for Exadata database servers as shown in the slide.
- For Oracle Data Guard, set the redo apply parallelism to 16 or less by using the following command:

```
SQL> RECOVER MANAGED STANDBY DATABASE … PARALLEL 16;
```

- As a general guide, the number of active database processes should not exceed five times the value of the CPU_COUNT parameter setting. If this ratio is exceeded, you should consider allocating additional CPU resource to the database. For very CPU-intensive applications, you may need to limit the number of active processes to the value of CPU_COUNT.
- Where it is feasible, use connection pools to manage application connections. This enables more efficient use of connections and minimizes resource usage associated with establishing and terminating connections.

- If the application tier cannot change, using the shared server architecture for many short requests and transactions can dramatically reduce the database process count. Check your application documentation for any restrictions or specific guidance about using the shared server architecture of Oracle Database.
- Consider using the connection rate limiter feature of Oracle Net Listener to limit the rate of new connections handled by the listener. This feature can be used to prevent potentially adverse performance impacts associated with a large number of simultaneous connections.

Other Recommendations

- Configure I/O Resource Management (IORM).
- Set `DB_RECOVERY_FILE_DEST_SIZE`:
  - sum(`DB_RECOVERY_FILE_DEST_SIZE`) < usable FRA size

The following additional recommendations apply for database consolidation:

- Exadata IORM provides powerful facilities that are very useful for database consolidation. IORM is covered in detail in the lesson titled “I/O Resource Management.”
- Ensure that the `DB_RECOVERY_FILE_DEST_SIZE` parameter for each database is set in such a way that the sum of all the sizes is smaller than the usable FRA area. The usable FRA area is the size of the FRA less any space consumed by files such as the mirrored copies of the online redo log files and the control files.

Isolating Management Roles

- Separate DBA and ASM administration roles.
  - They can be configured as part of the OneCommand process.
- Separate `OSDBA` and `OSOPER` groups, and separate administrator OS accounts for each database.
  - They must be configured manually.
  - They can be used for audit purposes or to enforce isolation.
  - Enforced isolation requires a separate Oracle Database software installation for each database.
- Separate disk groups for each database.
  - See the discussion on non-standard storage configurations earlier in this lesson.
  - Database-scoped Exadata storage security can also be implemented to enforce storage isolation.

A common requirement in a consolidation setting is the ability to isolate the management functions associated with each database. With Exadata, several options exist to define and enforce isolation at various levels.

- During the initial installation and configuration process, there is an option to create separate database administration and ASM administration accounts. Using this option, an operating system account, named `grid` by default, is created to manage ASM and Grid Infrastructure. Another operating system account, named `oracle` by default, is created with the privileges required to administer Oracle Database. Privileged groups are also created and associated with the `OSDBA`, `OSOPER`, `OSASM`, ASM `OSDBA`, and ASM `OSOPER` roles.
- In addition, customers can manually configure separate `OSDBA` and `OSOPER` groups, and separate OS user accounts for each database. Separate OS users can be used as a mechanism to audit the activity of different administrators. They can also be used to enforce management isolation between different databases. For example, a specific OS account could be given rights to manage a specific database but could be forbidden from even connecting to other databases. Enforced isolation requires a separate Oracle Database software installation for each database.

- Another level of isolation can be achieved by using different disk groups for each database. This can be used to achieve a physical separation so that each database is stored on physically separate disks. Some alternative storage configuration options were discussed earlier in this lesson. In addition to separate disk groups, database-scoped Exadata storage security can be used to enforce storage isolation. Using this feature, different Exadata grid disks can be associated with different databases. For a description of how to implement Exadata storage security, see the lesson titled “Exadata Storage Server Configuration.”

Schema Consolidation Recommendations

- Isolate application schemas into their own tablespaces:
  - Simplifies space management
  - Enables fast and granular recovery using tablespace point-in-time recovery (TSPITR)
- Tune tablespace backup, restore, and recovery practices to minimize disruption between schemas.
  - Understand the overheads and requirements of different approaches:
    - TSPITR
    - Flashback technologies
    - Logical export and import
    - Others
- Consider using image copies for faster TSPITR.

For schema-level consolidation, the main consideration is whether or not the application schemas are suitable to co-exist within the same database. If schema-level consolidation is feasible, the result is a single database, which should be configured and managed in line with all the standard recommendations for Exadata. In addition, the following recommendations and considerations apply:

- Each application schema should be contained in a separate tablespace or set of tablespaces. This enables easy and efficient space management. It also facilitates the use of tablespace point-in-time recovery (TSPITR) to perform recovery of a specific application schema without impacting the other tablespaces and schemas.
- Tune your backup and recovery practices so that individual application schemas can be maintained with minimal disruption to other applications. This includes understanding the overheads and requirements of different approaches, and ensuring that the required system resources are available to support the desired objectives.
- If TSPITR is considered, also consider whether image copies can be used to enable faster recovery. In this case, RMAN does not need to restore the data files from a backup.

Refer to My Oracle Support note 1386048.1 for a detailed discussion of schema recovery options in a consolidated environment.

Consolidation Using Virtual Machines

- OVM is supported on Exadata database servers:
  - Run multiple VMs on each database server
  - Share Exadata and maintain strict isolation:
    - Host multiple databases, customers, services and so on
    - Isolation for databases with strict certification requirements
  - Single Root IO Virtualization (SR-IOV) delivers near-raw InfiniBand networking performance
  - Specify CPU and memory usage
  - License software at the VM level
- Recommendation for consolidation:
  - Use normal database consolidation in general
  - Use VMs where they add specific benefits
  - Avoid one VM for each database; overhead is too high

ORACLE

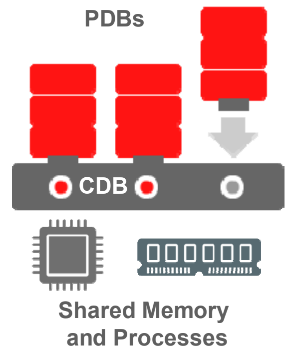
Commencing with Exadata software release 12,1,2,1.0, you have the option of running Oracle Virtual Machine (OVM) on the Exadata database servers. Using this option you can run multiple VMs on each database server. This configuration enables multiple databases to share Exadata while maintaining a very high degree of isolation between them. This is a very desirable feature in various shared-service scenarios, and also in cases where applications have strict certification requirements, which demand specific software versions.

To ensure uncompromised performance, Exadata VMs use high speed InfiniBand networking with Single Root IO Virtualization (SR-IOV). This delivers performance within a VM that is similar to raw hardware performance. Exadata VMs can dynamically expand or shrink CPU and memory usage based on workload requirements. Exadata VMs are considered Trusted Partitions and therefore software can be licensed at the virtual machine level instead of the physical processor level.

Oracle recommends the judicious use of VMs for consolidation on Exadata. In general, you should continue to use the recommendations for consolidating multiple databases and only use virtual machines to address specific requirements that normal database consolidation cannot. Using a separate VM for every database is not recommended because of the substantial overhead of running many copies of the operating system on each server.

Consolidation Using Oracle Multitenant Architecture

- Oracle Multitenant Architecture:
  - One Container Database (CDB)
    - One set of background processes
    - One System Global Area
  - Multiple Pluggable Databases (PDBs)
    - Up to 252 PDBs in one CDB
- Benefits of other consolidation approaches without associated drawbacks:
  - PDBs behave like independent databases
  - Use fewer machine resources
  - Avoid schema clashes
- Highly recommended for Exadata:
  - Implement in accordance with Exadata best-practices
  - Configure DBRM and IORM
  - Implement without VMs for best efficiency



ORACLE

Introduced in Oracle Database 12*c* Release 1, Oracle Multitenant delivers a new architecture that allows a multitenant container database (CDB) to host many pluggable databases (PDBs). Existing databases can be easily converted to a PDB and consolidation is achieved by plugging in multiple PDBs into a single CDB. In Oracle Database 12*c* Release 1, a CDB can contain up to 252 PDBs.

The multitenant architecture combines the benefits of other consolidation approaches, while avoiding their drawbacks. Specifically:

- A CDB has a single set of background processes and shared memory area (SGA) that is shared by all of the PDBs. This architecture requires less CPU and memory compared to traditional approaches of consolidating multiple independent databases onto a single physical machine, or using multiple VMs.
- The PDBs are isolated from each other just like independent databases. Consequently, the application schema clashes that are commonly encountered with schema-based consolidation are completely avoided.

For these reasons, the Oracle multitenant architecture is highly recommended for database consolidation on Exadata.

Implementation should be performed in accordance with the best-practices for Exadata, and you should use Oracle database resource management (DBRM) in conjunction with Exadata I/O Resource Management (IORM) to ensure that your PDBs share Exadata resources effectively.

While a CDB can be deployed in either physical or virtual environments, the highest management and performance efficiency is achieved with physical implementations. The CDB itself becomes the virtualization technology for the database tier, eliminating the overhead of multiple database instances, VMs and guest operating systems.

General Maintenance Considerations

- Additional headroom and specialized procedures may be required to accommodate server outages or rolling patches.
  - How is the environment impacted by the absence of a database server or an Exadata Storage Server?
  - Use policy-managed databases and services to adapt.
- Isolate changes in a new Oracle home directory, and then migrate and validate each database.
  - Validate in test and standby first if possible.
- Maintain only a few active Oracle software versions:
  - Example: 11.2.0.1 BP9, 11.2.0.2 BP7, and 11.2.0.3 BP2
  - Remember that the Grid Infrastructure version must be the same as, or later than, the database kernel version.

Maintaining a consolidation platform may require additional headroom or the adoption of customized procedures. For example, the temporary absence of a storage cell during a rolling patch will have a greater impact on a database that uses only a few cells compared with a database that is striped across all 14 cells in a Full Rack Database Machine. Consider how the absence of a server might impact the environment and plan accordingly. Also use features such as policy-managed databases and database services so that the environment can more easily absorb the impact of a server outage.

When updating Oracle Database software versions, it is generally good practice to isolate the update in a new Oracle home directory. This is particularly applicable in a consolidation setting because it allows databases to be migrated and validated separately.

Remember to manage the collection of Oracle software versions and maintain only a few active versions because a proliferation of different versions will likely cause additional maintenance complexity. Also, remember that the Grid Infrastructure software version must be the same as, or later than, the latest active database kernel version. In other words, the Grid Infrastructure software version cannot be earlier than any of the database kernel versions.

Quiz

Exadata is suitable for:

a. Application schema consolidation
b. Database consolidation
c. Both of the above

ORACLE

Answer: c

Quiz

When practicing database consolidation on Exadata, Oracle recommends creating a separate set of ASM disk groups for each database.

a. True
b. False

Answer: b

The recommended storage configuration for consolidation is to use one shared ASM storage grid for each consolidated environment. The default disk group configuration resulting from the installation and configuration process should be used. The recommended storage configuration has the following major benefits. Firstly, it is conceptually simple and inherently easy to manage, it is also the most familiar and supportable configuration. Secondly, it is a balanced configuration which is suitable for most circumstances and provides high levels of availability and performance.

Although it is possible to create a separate set of disk groups for each database, and although such a configuration does provide specific benefits in some circumstances, Oracle does not recommend a separate set of ASM disk groups for each database. Customers should consider the benefits and limitations of using multiple disk groups sets before implementing such a configuration.

Summary

In this lesson, you should have learned how to:

- Describe the options for consolidating multiple databases on Exadata
- Explain the benefits and costs associated with different options

Additional Resources

- Documentation and white papers
  - Best Practices For Database Consolidation On Oracle Exadata Database Machine
    http://www.oracle.com/technetwork/database/features/availability/exadata-consolidation-522500.pdf

11

Migrating Databases to Exadata Database Machine

ORACLE

Objectives

After completing this lesson, you should be able to:

- Describe the general steps to migrate your database to Exadata
- List the main approaches for migrating your database to Exadata
- Identify the most appropriate approach for different circumstances

ORACLE

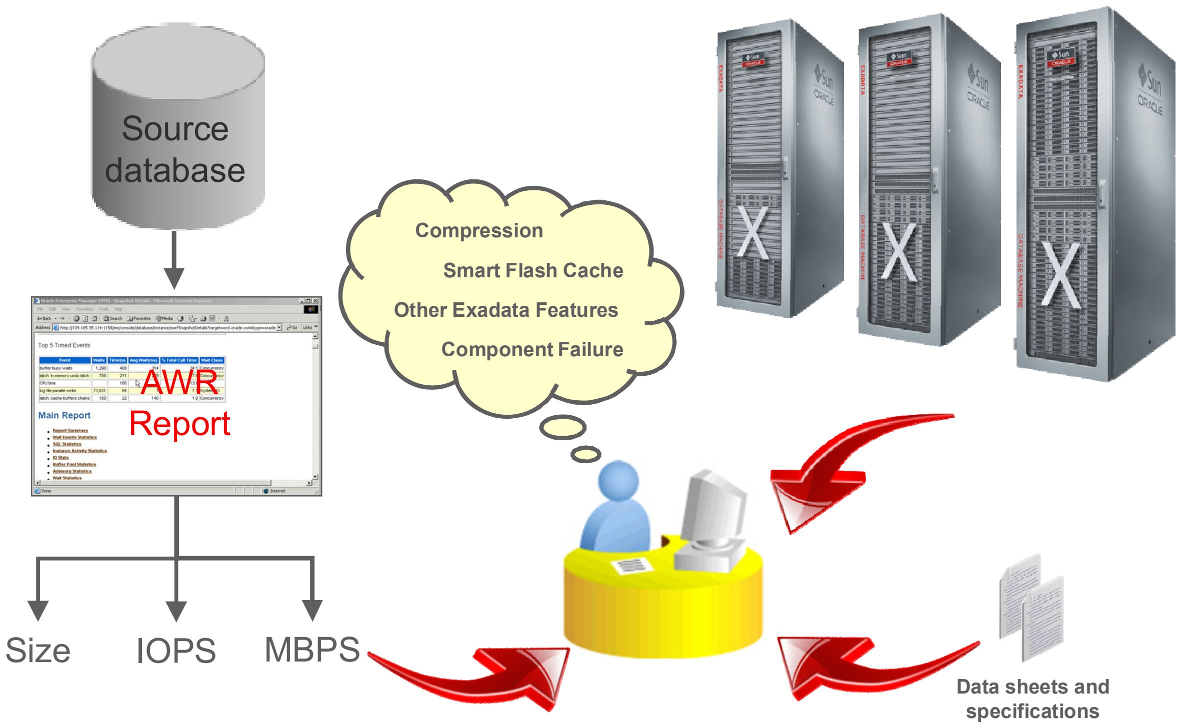
Migration Best Practices: Overview

Original storage to Exadata

1. Capacity planning
2. Choosing the migration approach
3. Migration
4. Post-migration tasks

The slide depicts the four important phases to migrate your existing databases to Exadata. Each phase is discussed in the remainder of this lesson.

Performing Capacity Planning



The biggest Exadata capacity planning challenge is understanding the difference between your existing storage and Exadata storage. To determine the storage requirements, you must understand the I/O characteristics of your current environment. Collect the size and throughput of your current system. The key measures of I/O throughput are I/Os per second (IOPS) and megabytes per second (MBPS). Use the system statistic *physical I/O disk bytes* to derive the current MBPS for the system. Use the system statistics *physical reads* and *physical writes* to determine the current IOPS for the system. These statistics are available in an Automatic Workload Repository (AWR) report.

After you understand the capacity of your current system, you can determine the appropriate Exadata storage configuration. It is important to size both for performance and capacity. Use the Exadata performance and capacity metrics published at http://www.oracle.com/exadata to assist with sizing.

It may be possible to achieve much greater effective throughput by making efficient use of Exadata Smart Flash Cache and Hybrid Columnar Compression. The precise impact of these technologies should be verified by testing.

It is also important to consider failures when planning capacity. Exadata cell and disk failures are transparently tolerated using Automatic Storage Management (ASM) redundancy. However, it is best practice to ensure that post-failure I/O capacity is sufficient to meet your redundancy requirements and performance service levels.

Exadata Migration Considerations

- Platform is 64-bit Intel.
  - Byte order is little-endian.
- Exadata, ASM, and database software must run release 11.2.0.1 or greater.
- ASM disk group attributes:
  - `COMPATIBLE.ASM=11.2.0.0.0` (or later)
  - `COMPATIBLE.RDBMS=11.2.0.0.0` (or later)
  - `CELL.SMART_SCAN_CAPABLE=TRUE`
  - `AU_SIZE=4M`
- Database initialization parameters:
  - `COMPATIBLE=11.2.0.0.0` (or later)
  - `DB_BLOCK_CHECKSUM=TYPICAL|FULL`
- Database extent sizes should be a multiple of 4 MB.

ORACLE

The following are the configuration requirements and recommendations for Exadata that you must know while planning your database migration:

- Exadata is a 64-bit Intel-based platform. The byte order is little endian.
- The Exadata Storage Server, ASM, and database software on Exadata must be release 11.2.0.1 or later.
- For ASM disk groups containing Exadata grid disks, specify the attribute settings shown in the slide. Although these are not mandatory, they are highly recommended to enable all of the Exadata Storage Server features and optimize performance.
- To enable all the Exadata features in your database, the Oracle Database software must be release 11.2.0.1 or later, and you must set the `COMPATIBLE` initialization parameter as shown in the slide. Because Exadata storage is Hardware Assisted Resilient Data (HARD) compliant, you should set `DB_BLOCK_CHECKSUM` to `TYPICAL` or `FULL` to ensure that checksums are computed and stored in the blocks. Note that this initialization parameter is set to `TYPICAL` by default.
- Exadata performs best when scanning contiguous chunks at least 4 MB in size. To ensure that this occurs, the ASM disk group allocation unit (AU) size (`AU_SIZE`) must be set to 4 MB and the database extent sizes should be a multiple of 4 MB.

Note: The parameters highlighted in red are required to enable smart storage capabilities.

Choosing the Right Migration Path

- Determine what to migrate:
  - Avoid methods that migrate what you discard.
- Consider the configuration of the source system:
  - Source Oracle Database version and platform matters.
  - Target system is fixed: 11.2 or later, ASM, and little-endian.
- Weigh up the costs and favor methods that facilitate best practices:
  - Implementing best practices is important in the long term because your future performance depends on it.
  - Remember:
    - ASM AU size of 4 MB can be set only at disk group creation.
    - Database extent sizes are set at extent allocation.

ORACLE

You can migrate an existing database to Exadata by using several different approaches. Typical approaches are discussed later in this lesson. Before considering them, it is worthwhile to consider some factors that can help in choosing the right migration path.

Before choosing a migration approach, you must clearly define what you want to migrate. This will help you to avoid wasted effort, such as migrating data that is not required. Clearly defining the scope of the migration also helps you to identify the source systems. You need to understand the source systems because their composition may limit the available migration options. For example, Exadata is a little-endian platform, so if you are migrating from a big-endian platform, some physical migration approaches are not feasible or require extra processing. Also, the use of database features, such as materialized views or object data types, may impose restrictions on some migration methods.

You must also weigh up the short term requirement to perform the migration with the long term impact of using the selected migration approach. Specifically, you may need to rule out what seems to be an easy and convenient migration approach if the resulting database configuration is sub-optimal or compromised in some way. For example, Exadata performs best with an ASM AU size of 4 MB and database extents that are a multiple of 4 MB. If the source database extent sizes are not a multiple of 4 MB and it is impractical to reorganize the database before migration, then you might favor a migration approach that allows you to reorganize the database during the migration. If you choose an approach that does not allow the extents to be reorganized, you may be able to deliver a quicker and easier migration. However, you may also end up paying an ongoing performance penalty.

Logical Migration Approaches

Migration approach	Cross-version support	Supports physical reorganization	Cross-platform support	Complexity	Migration outage time
Logical standby database *Minimal down time*	Limited, only for rolling upgrade	Yes, some limitations	Limited, little-endian only	High	Low
Insert-as-select *Manual migration using SQL*	Yes	Yes	Broad	Depends on schema complexity	Depends on data volume
Data Pump Export/Import *For migration during a planned maintenance period*	Yes	Yes	Broad	Low	Depends on data volume

ORACLE

Migration to Exadata can be achieved in many different ways. In the next part of this lesson, different approaches are introduced. Because every migration is subtly different, no detailed migration procedure is presented. Rather, different approaches are outlined along with their key characteristics and restrictions.

The table in the slide summarizes some logical migration approaches by using technologies that are embedded into Oracle Database. Using a logical approach, you can change the database extent size and alter other physical characteristics of your database, such as the database character set, which is not possible by using a physical migration approach.

- **Logical standby database:** If your application service-level agreements permit little or no down time, you can use an Oracle Data Guard logical standby database to replicate the database, and track and merge the changes while the source database runs. After the configuration is established, the logical standby database (on Exadata) can be switched to assume the role of primary database and the original source database can be decommissioned.

  The procedure documented in My Oracle Support note 737460.1 can be used to change some physical storage attributes, such as segment extent sizes. Refer to My Oracle Support note 1085687.1 for information about heterogeneous platform support (and associated limitations) in the same Data Guard configuration. Support for different source and target database versions is only provided for rolling upgrades.

It is worth noting that there are some limitations that contradict the use of logical standby databases. For example, logical standby databases are unable to natively handle all Oracle Database data types. Although there are methods to overcome this limitation, the extra effort required to implement and maintain these methods should not be overlooked. Also, a logical standby database will fail to duplicate NOLOGGING operations that are conducted on the primary database. For details, see *Oracle Data Guard Concepts and Administration*.

- **Insert-as-select:** You can migrate a manageable number of tables manually using SQL. Using a variety of different methods (manual scripting, Data Pump, SQL*Developer, and so on) you can clone the structure of your tables, making necessary adjustment to physical storage attributes along the way. Then, you can use an `INSERT ... AS SELECT` SQL command in conjunction with a database link to copy the data for each table. You can contain the code for multiple tables in SQL scripts or you can encapsulate code in PL/SQL routines. This method is preferred in situations where users can easily repurpose existing code that copies data between databases.
- **Data Pump:** If a suitable maintenance window can be accommodated and if the size of the database is not prohibitively large, use Data Pump to move the data in bulk from the source system to Exadata. Data Pump is easy to use and provides broad support across different platforms and database versions.

Physical Migration Approaches

Migration approach	Cross-version support	Supports physical reorganization	Cross-platform support	Complexity	Migration outage time
RMAN Backup/Recovery *Backup-based migration*	No	No	Linux x86-64 only	Low to medium	Depends on data volume
Data Guard physical standby *Minimal down time*	No	No	Limited, little-endian only	Medium	Low
Transportable Database *Migration to a different platform with the same endian format*	No	No	Little-endian only	Low to medium	Depends on data volume
Transportable Tablespaces *Migration to a different platform with a different endian format*	Yes	No	Broad	Low to medium	Depends on data volume
Unplugging and plugging, or remote cloning, a PDB or Non-CDB *Requires 12c or later source*	Limited	No	Little-endian only	Low	Depends on data volume

The table in the slide summarizes some physical migration approaches by using technologies that are embedded into Oracle Database:

- **RMAN Backup/Recovery:** Oracle Recovery Manager (RMAN) is an Oracle Database client that performs backup and recovery tasks on Oracle databases. You can use RMAN to migrate data simply by transferring a backup of your source database to the target environment and restoring it there. If your source database resides on Linux x86–64 (like Exadata), and it uses Oracle Database 11*g* Release 2 or Oracle Database 12*c* Release 1, you can use RMAN to restore a backup of your source database on Exadata. RMAN also provides an active database duplication feature, which performs duplication over a network link between the source and target databases. You must consider the size of your source database, and the speed and reliability of your network connection to determine the feasibility of this approach. For information about using RMAN, refer to the *Oracle Database Backup and Recovery User's Guide*.

  Note: RMAN is also used to facilitate some of the other migration methods listed in the slide.

- **Data Guard physical standby:** An Oracle Data Guard physical standby database is a block-for-block replica of a primary database. You can use Data Guard to replicate your source database to Exadata. Afterwards you can decouple the databases and use the physical standby database as your new master.

  You can use this method in conjunction with source databases from a selection of little-endian platforms. See My Oracle Support note 413484.1 for information about heterogeneous platform support. Also, the primary and standby databases must have the same compatibility setting, which means that your source database must be upgraded to a version of Oracle Database supported by Exadata before Data Guard is configured.

  For details, see *Creating a Physical Standby Database* and *Creating a Standby Database with Recovery Manager* in *Oracle Data Guard Concepts and Administration*.
- **Transportable Database:** This method works in conjunction with RMAN to migrate whole databases between platforms that share the same endian format. The result is a block-for-block replica of the source database. Consequently, the transportable database method is useful in cases where it is not necessary to physically reorganize the source database.

  Though conceptually similar, the transportable database method is substantially different from transportable tablespaces. The transportable database method involves copying an entire database, including the `SYSTEM` tablespace, from one platform to another. Because the whole database is copied, containment checks are unnecessary and no Data Pump export and import is required. RMAN is used to perform the required backup, conversion, and restoration operations.

  The transportable database method only works across platforms that share the same endian format. Therefore, your source database must reside on a little-endian platform in order facilitate transport to Exadata.

  For further details, see *Transporting Data Across Platforms* in *Oracle Database Backup and Recovery User's Guide*.
- **Transportable Tablespaces:** Use the Transportable Tablespaces feature to migrate tablespaces from your current system to a new database hosted on Exadata. This is the only physical database migration method that provides broad cross-platform support and supports migration from earlier Oracle Database versions.

  Using this feature, data is migrated by copying the source data files to the target system, and Data Pump is used to migrate the associated data dictionary definitions. If you are migrating from a big-endian platform to Exadata (little-endian), extra processing is required to perform a conversion.

  Administrative tablespaces, such as `SYSTEM` and `SYSAUX`, cannot be transportable tablespaces.

  For further details, see *Transporting Tablespaces Between Databases* in *Oracle Database Administrator's Guide*.

- **Unplugging and plugging, or remote cloning, a PDB or Non-CDB:** You can migrate an Oracle Database 12*c* pluggable database (PDB) to Exadata by unplugging the PDB from the source container database (CDB) and plugging it into a CDB on Exadata. Likewise, you can migrate an Oracle Database 12*c* non-CDB to Exadata by plugging the non-CDB into a CDB on Exadata.

  Cloning a Remote PDB or Non-CDB is very similar to unplugging and plugging in a PDB or plugging in a Non-CDB. The major difference is that remote cloning uses a database link to transfer the data as part of running the `CREATE PLUGGABLE DATABASE` command. As a result, remote cloning is even simpler than preparing, transporting and plugging in a PDB. However, since remote cloning depends on transporting the data over a database link, you must consider the size of your source database and the speed of your network connection to determine the feasibility of this approach.

  Approaches using PDBs are very attractive because of their simplicity. However, the specific requirements for PDBs make these approaches suitable in fewer situations than other methods, such as transportable tablespaces. The requirements include:

  - The source and target database version must be 12.1 or later. Furthermore, if you are creating a PDB by cloning a non-CDB, then both the target CDB and the source non-CDB must be running Oracle Database 12*c* version 12.1.0.2 or later.
  - The source and target platform must have the same endian format, which is little-endian for Exadata.
  - Ideally, the source and target CDBs must use the same character set, and Oracle recommends using the AL32UTF8 character set for all new databases on Exadata.

    Note: For CDBs that use the AL32UTF8 character set, you can have PDBs that use a multibyte character set that is a binary subset of AL32UTF8, such as UTF8. However, complications may arise if the different character sets have different maximum character widths.

  For more information, see *Creating a PDB by Plugging an Unplugged PDB into a CDB*, *Creating a PDB Using a Non-CDB*, and *Cloning a Remote PDB or Non-CDB* in *Oracle Database Administrator's Guide*.

Reducing Down Time for Migration by Using Transportable Tablespaces

Traditional Migration Using Transportable Tablespaces	Migration Using Cross Platform Incremental Backup
1. Make the source tablespaces READ ONLY. 2. Transfer the data files to the destination system. 3. Convert the data files to the destination system endian format. 4. Export the metadata of objects in the tablespaces from the source database by using Data Pump. 5. Import the metadata into the destination database. 6. Make the destination tablespace READ WRITE. **Data unavailable for update during entire process.**	Phase 1 - Prepare Phase 1. Transfer data file copies to the destination system. 2. Convert data file copies to the destination endian format. Phase 2 - Roll Forward Phase 1. Create and incremental backup on the source system. 2. Transfer the incremental backup to destination system. 3. Convert the incremental backup to the destination system endian format and apply it to the destination data file copies. Repeat Phase 2 until the destination system is almost up to date. **Source database remains available for updates throughout Phase 1 and Phase 2.** Phase 3 - Transport Phase 1. Make the tablespaces in the source database READ ONLY. 2. Repeat Phase 2 once more to bring the destination data files up to date with the source data files. 3. Export the metadata of objects in the tablespaces from the source database by using Data Pump. 4. Import the metadata into the destination database. 5. Make the tablespaces in the destination database READ WRITE. **Data unavailable for update in Phase 3 only.**

As mentioned in the preceding slide, Transportable Tablespaces facilitate the physical migration approach that provides the broadest support for different source platforms and database versions. When you use Transportable Tablespaces to migrate data between systems that have different endian formats, the amount of down time is directly proportional to the size of the data set being moved.

Oracle has released a capability called Cross Platform Incremental Backup, which when used in conjunction with Transportable Tablespaces, significantly reduces the amount of down time required to migrate data to Exadata.

The slide shows an outline of the new process in comparison with traditional migration using Transportable Tablespaces. For complete details about this process, refer to My Oracle Support note 1389592.1.

Note that Cross Platform Incremental Backup does not affect the amount of time it takes to perform other migration actions such as metadata export and import. Therefore, databases that have very large amounts of metadata will see limited benefit if the migration time is dominated by metadata operations, not data file transfer and conversion.

Note: Cross Platform Incremental Backup requires Exadata running Linux with at least Oracle Database version 11.2.0.2 and Bundle Patch 12 for Exadata installed.

Other Approaches

- Data Integration Tools
  - Oracle GoldenGate can be used to transform and migrate data while the source systems continue to run.
- Hybrid Approaches
  - For example, use Transportable Tablespaces to migrate data from the current production database to a staging database outside Exadata, and then use Data Pump to unload data from the staging database and load it into Exadata.

The preceding slides introduced several migration approaches based on tools and technologies that are part of Oracle Database. In addition, there are many other migration technologies and approaches:

- Data integration technologies such as Oracle GoldenGate can be used to migrate, and if necessary transform data. Typically, such environments provide a high degree of flexibility, along with a productive way of defining and implementing the required data integrations and transformations. Also, these technologies are often used to facilitate data migration while the source system continues to run.
- There are also many potential hybrid approaches that combine two or more of the methods introduced in this lesson. For example, your situation might determine that Transportable Tablespaces are a convenient way to migrate data to Exadata. However, the physical organization of the data in the Transportable Tablespaces may not be ideal for Exadata, so you may choose redefine the tables by using a series of `CREATE ... AS SELECT SQL` commands, or to reload the data into fresh segments using Data Pump.

Post-Migration: Best Practices

- Check whether ASM disk groups are balanced:
  - Manual rebalance may be needed.
  - Script is available in My Oracle Support note 367445.1.
  - Enterprise Manager alert for disk group imbalance
- Assess index requirements:
  - With Exadata, full scans might deliver acceptable performance.
  - Determine the indexes that are not required and remove them.
- Configure I/O Resource Management.

ORACLE

Consider the following after migrating databases to Exadata:

- One of ASM's core functions is to ensure that data is evenly distributed across all disks in a disk group. This happens automatically. However, occasionally, a disk group may become imbalanced due to uncommon errors, such as a failed rebalance. It is, therefore, an operational best practice to check disk groups on a regular basis and run a manual rebalance if necessary. A script is available to check disk group balance in My Oracle Support note 367445.1. Also, Enterprise Manager Grid Control displays an alert if a disk group becomes unbalanced beyond a customizable threshold.
- The superior scan rates available on Exadata make it possible that indexes, previously required for good performance, are no longer required. You should assess execution plans that use indexes to see if they would run acceptably with Smart Scans. To determine if queries would perform acceptably without an index, you can make the index invisible to the optimizer. An invisible index is maintained by DML operations, but it is not used by the optimizer for queries. To make an index invisible, use the following command:

```
ALTER INDEX <index_name> INVISIBLE;
```

- After you perform the preceding tasks, you can configure I/O Resource Management (IORM).

Quiz

What are two recommended practices when migrating to Exadata storage?

a. Double the size of your SGA.
b. Configure ASM allocation units to 4 MB.
c. Double the size of your ASM disk groups.
d. Configure database extents to multiples of 4 MB.

ORACLE

Answer: b, d

Quiz

Which of the following physical migration approaches is most universally applicable?

a. Transportable Database
b. Data Guard Physical Standby Database
c. Transportable Tablespaces

Answer: c

Using Transportable Tablespaces, you can migrate to Exadata from a much broader range of platforms than either of the other approaches.

Summary

In this lesson, you should have learned how to:

- Describe the general steps to migrate your database to Exadata
- List the main approaches for migrating your database to Exadata
- Identify the most appropriate approach for different circumstances

Additional Resources

- My Oracle Support notes
  - Changing Storage Definition in a Logical Standby Database
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=737460.1
  - Data Guard Support for Heterogeneous Primary and Standby Systems in Same Data Guard Configuration
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=413484.1
  - Script to Report the Percentage of Imbalance in all Mounted Diskgroups
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=367445.1

ORACLE

Practice 11 Overview: Migrating to Exadata Using Transportable Tablespaces

In this practice, you will use Oracle Recovery Manager (RMAN) in conjunction with the Transportable Tablespaces feature of Oracle Database to migrate data from a big-endian platform to Exadata (which is a little-endian platform).

12

Bulk Data Loading

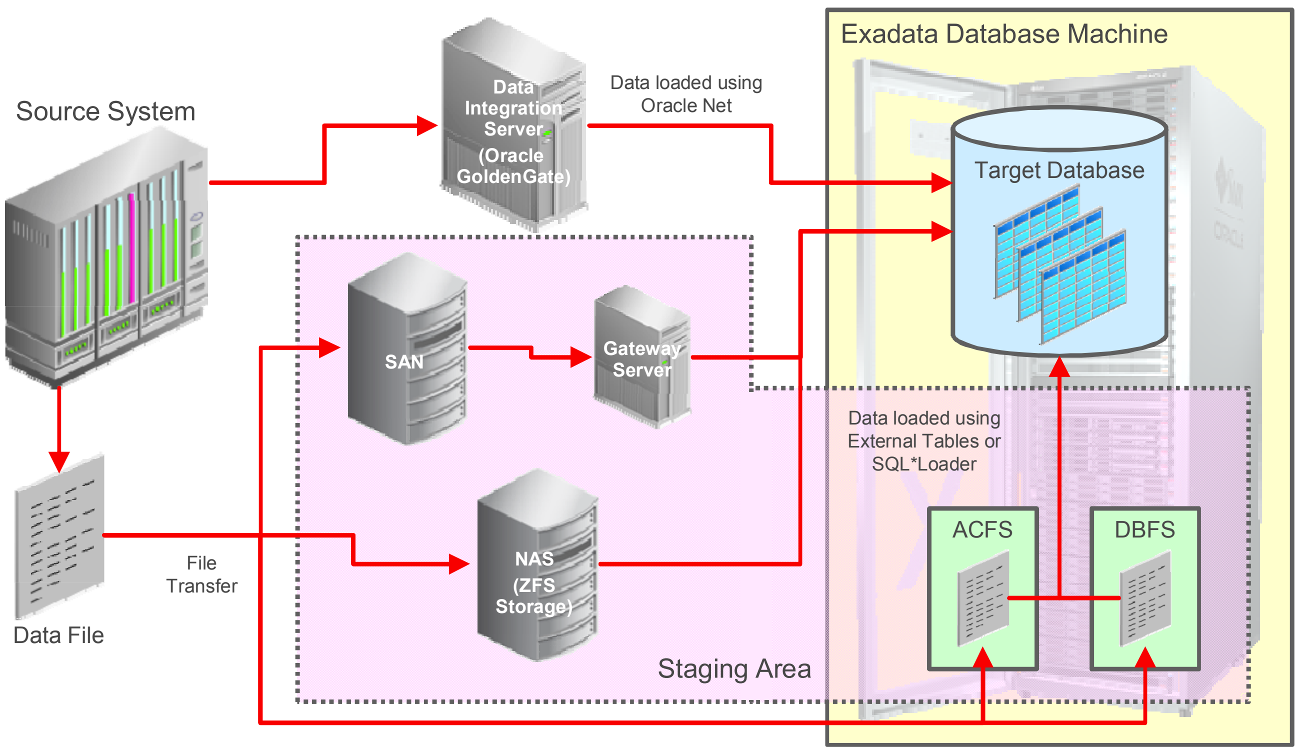
Objectives

After completing this lesson, you should be able to:

- Describe alternative architectures for bulk data loading into Exadata Exadata
- Configure the Database File System (DBFS) feature for staging input data files
- Use external tables to perform high-performance data loads

ORACLE

Bulk Data Loading Architectures for Exadata



The slide illustrates common architectures for bulk data loading into Exadata. These include:

- Loading through a data integration server such as Oracle GoldenGate.

  Using this approach data is harvested from the source system and loaded directly into the Exadata database. The integration server may process or aggregate the data along the way. Specific implementation details vary depending on the integration server platform being used.
- Loading file-based data using a staging area.

  Using this approach, data from the source systems is written into files, which are transported to a staging area that is accessible to Exadata. Files can be transported to the staging area over a network, or by physically transporting file storage media such as disk, tape or flash storage. Data from the staging area can be loaded by various means; however, external tables and SQL*Loader are the most commonly used mechanisms.

  The staging area can be hosted on Exadata using ASM Cluster File System (ACFS) or Database File System (DBFS). Using the available space on internal database server disk drives for staging data files is highly discouraged. Alternatively, the staging area can also be hosted on storage infrastructure separate from Exadata, such as SAN or NAS. These options are outlined next in this lesson.

Staging Data Files Using DBFS

DBFS enables the database to be used as a file system:

- DBFS provides shared storage for staging or storing data files, scripts, reports, and other application files.
- Files are stored as SecureFiles LOBs in database tables that are stored in Exadata.
- Files are exposed to the database servers by using a POSIX-compatible file system interface.
- Files are transactionally consistent with database data.
- Files can leverage Oracle Database availability features:
  – ASM mirroring, RMAN, Data Guard, Flashback, and so on.

ORACLE

DBFS is an Oracle Database feature that enables the database to be used as a high-performance, POSIX-compatible file system.

In DBFS, files are stored as SecureFiles LOBs. A set of PL/SQL procedures implement the file system access primitives, such as open, close, create, and so on. The `dbfs_client` utility enables a DBFS file system to be mounted on the Exadata database servers. It also provides the mapping from file system operations to database operations.

Because DBFS stores files inside an Oracle database, the files are transactionally consistent with other data in the database. Files can also take advantage of the high-availability features of Oracle Database on Exadata. Files can be protected by using ASM mirroring, backed up and recovered using RMAN, replicated for disaster protection using RMAN, and managed over time using flashback technologies.

Staging Data Files Using ACFS

ASM Cluster File System (ACFS):

- Fully-featured cluster file system built on ASM:
  - Snapshots
  - Replication
  - Advanced auditing
  - Realm-based security and encryption
  - Tagging API for metadata
- Comparable performance with DBFS for bulk loads
- Separate from all databases:
  - Files are stored outside the Oracle databases
  - Files can only be accessed through file system APIs
  - Cannot leverage database features such as RMAN and Data Guard
- Requires Oracle Grid Infrastructure version 12.1.0.2 or later
  - Otherwise, no Exadata-specific implementation requirements

ORACLE

You may be able to use ACFS as an alternative to using DBFS.

ACFS is a file system that is built directly over Oracle Automatic Storage Management (ASM). ACFS stores files directly inside ASM, without using an Oracle database. ACFS provides richer file system capabilities and performs better than DBFS for creating and deleting large numbers of small files; however, performance is comparable for bulk data loads.

Because ACFS files are not stored in an Oracle database, you do not have automatic transactional consistency between your files and other database data. Also, you must manage ACFS separately from your Oracle databases. For example, you will need to use a separate file-based backup and recovery system for ACFS.

To use ACFS on Exadata requires Oracle Grid Infrastructure version 12.1.0.2 or later. Otherwise, there are no Exadata-specific implementation requirements for ACFS on Exadata. For more information on ACFS, see the *Oracle Automatic Storage Management Administrator's Guide*.

Staging Data Files Using External File Systems

- Oracle Sun ZFS storage provides a high-performance and cost-effective alternative that:
  - Is tested and validated with Exadata
  - Can connect directly to the Exadata InfiniBand network
- Other NAS and NFS-based options are feasible.
  - Thorough testing is required to ensure acceptable performance.
- Connection to a SAN requires an intermediate server.
  - Exadata does not support direct SAN connection.
  - Performance may be limited by the intermediate server.

Oracle Sun ZFS Storage Appliance is an enterprise-class network-attached storage (NAS) environment, which can be connected directly to the Exadata InfiniBand network and can be used as a staging location for bulk data loads into Exadata. For more information, see the Oracle Sun ZFS Storage Appliance home page at https://www.oracle.com/storage/nas/index.html.

Other options using NAS and Network File System (NFS) technologies are also feasible. However, users should thoroughly test the proposed solutions to ensure that their performance is adequate.

Users wishing to stage data files using Storage Area Network (SAN) technologies must factor in the requirement for an intermediate server to act as an iSCSI or NFS gateway between Exadata and the SAN storage. An intermediate server is required because adding Fibre Channel adapters to the Exadata servers is not supported. Users should also note that the intermediate server may limit the throughput that can be achieved by using a SAN.

Comparison of Staging Area Options

Staging Area Technology	Potential Strengths	Potential Drawbacks
Database File System (DBFS)	• High performance data ingest with no network involvement • Transactional consistency with Oracle Database • Integration with Oracle Database for backup (RMAN) and disaster protection (Data Guard) • Works with Oracle Database 11*g*R2 and above	• Lack of advanced file system functionality, like snapshots • Files occupy space on Exadata
ASM Cluster File System (ACFS)	• High performance data ingest with no network involvement • Advanced file system functionality, like snapshots	• Lack of transactional consistency with Oracle Database • Lack of integration with Oracle Database for backup (RMAN) and disaster protection (Data Guard) • Files occupy space on Exadata • Requires Oracle Grid Infrastructure version 12.1.0.2

ORACLE

The slide contains a table that outlines the potential strengths and drawback of using either DBFS or ACFS as a staging area for bulk data loading into Exadata.

In essence, both options facilitate high performance data ingest because the staging area is maintained on Exadata. However, this comes at the cost of giving up space on your Exadata system to house the staging area. The main difference between the technologies are derived from the fact that DBFS stores data in an Oracle Database, while ACFS does not store data in an Oracle Database but offers richer file system functionality. Also, DBFS is supported with Oracle Database 11*g*R2 and above, while ACFS requires Oracle Grid Infrastructure version 12.1.0.2 as a minimum.

Consequently, the choice between these two options is usually determined by which Oracle Database software version you require, or whether the database integration in DBFS is more useful than the file system features of ACFS.

Comparison of Staging Area Options

Staging Area Technology	Potential Strengths	Potential Drawbacks
Network-Attached Storage (NAS) (including Oracle ZFS Storage Appliance)	• Files do not occupy space on Exadata • Comparatively cheap storage • Very large capacity • Advanced file system functionality • May be well integrated with source systems	• Additional infrastructure to procure and manage • Lack of transactional consistency with Oracle Database • Lack of integration with Oracle Database for backup (RMAN) and disaster protection (Data Guard) • Data ingest performance may be limited by network bandwidth, especially if not using InfiniBand
Storage Area Network (SAN)	• Same as NAS	• Same as NAS, plus: • Requirement for an additional gateway server to act as an NFS server, which may be a performance bottleneck

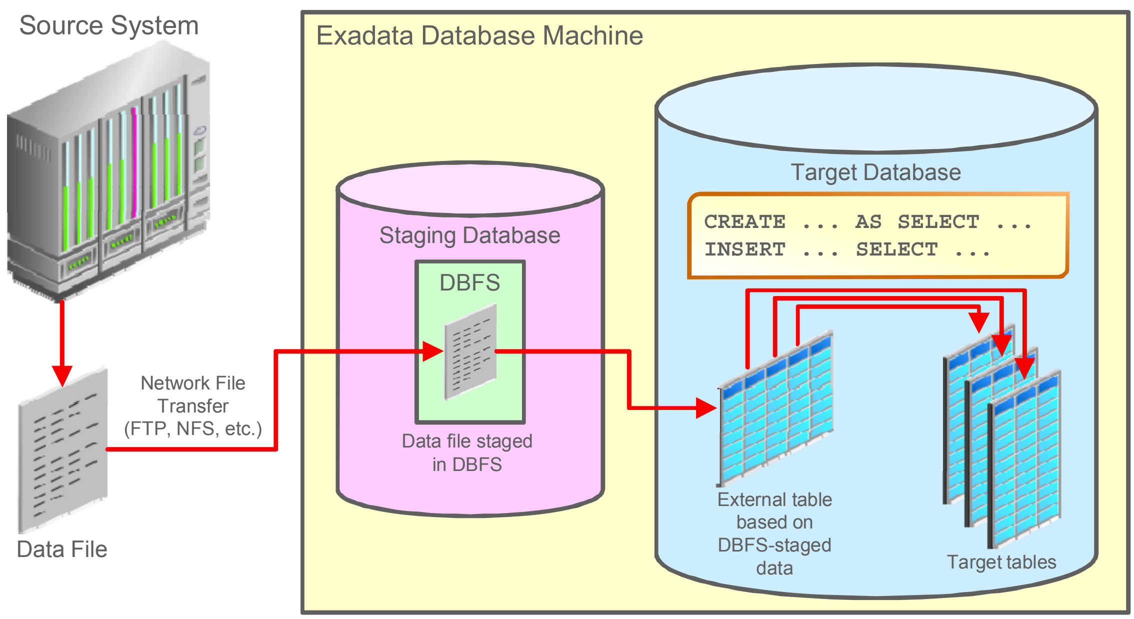
The slide contains a table that outlines the potential strengths and drawback of using NAS or SAN technologies to house the staging area for bulk data loading into Exadata.

Both options are well suited where there is an existing investment in NAS or SAN infrastructure, especially if it is already well integrated with the source systems.

Note that Oracle ZFS Storage Appliance is mentioned as an example of NAS storage that may be used. Oracle ZFS Storage Appliance supports connectivity to Exadata using the high-speed InfiniBand network. Consequently, the potential for network bandwidth limits to affect data ingest performance is inconsequential compared to a 1 Gbps Ethernet-based NAS system.

For the SAN-based option, you should factor in the requirement for an additional server to act as a gateway between Exadata an the SAN storage.

Bulk Data Loading Using Oracle DBFS: Overview



The slide illustrates the recommended approach for loading file-based data into Oracle by using DBFS. The remainder of this lesson describes this process in greater detail.

Note that DBFS can also be used as a generalized shared file system on Exadata. You can use the procedure for configuring DBFS (explained later in this lesson) to configure DBFS for purposes other than staging data files for bulk data loading.

Note also that the recommendations for preparing the data files and for configuring and loading the target database can be applied to loading file-based data into Oracle by using staging areas other than DBFS, such as ACFS or NFS.

Preparing the Data Files

- Data files can be in any format supported by the external table feature.
- To facilitate high-performance parallel access:
  - Oracle automatically divides the files into 10 MB granules.
    - Exceptions include compressed files, or serial devices.
  - If granules cannot be created:
    - Each file is treated as a granule
    - The number of files determines the maximum degree of parallelism
    - You must use multiple files to manually enable parallelism
  - General rules of thumb:
    - If using multiple files, try to keep them similar in size.
    - If the file sizes vary significantly, list them in order from largest to smallest in the external table definition.

The recommended approach for bulk data loading using Oracle DBFS relies on the external table feature of Oracle Database. The data files used for bulk data loading can be in any format supported by external tables. The process for creating the data files is outside the scope of this lesson and mostly depends on the facilities available in the source system.

Regardless of how the data files are created, the following should be taken into account to facilitate high-performance parallel access to the data files while they are being loaded:

- When accessing large data files through external tables Oracle automatically divides the files into 10 MB granules where possible. These granules can be processed in separate parallel processing threads. Oracle is unable to use this approach with compressed files or data read from a serial device such as a pipe or a tape device.
- If granules cannot be used, each separate data file can be treated as a granule and the number of files determines the maximum degree of parallelism that is available. You can manually divide a large file into separate smaller files and use them to manually enable parallelism.
- If you are using multiple input data files in conjunction with a single external table, you should try to keep the data files similar in size. If the file sizes do vary significantly, list them in order from largest to smallest in the external table definition.

Configuring a DBFS Staging Area

- DBFS can be configured in a separate staging database:
  - Use DBCA to create the staging database.
    - See My Oracle Support note 1191144.1 for recommendations.
- Create a bigfile tablespace for DBFS storage:

```
SQL> CREATE BIGFILE TABLESPACE DBFS DATAFILE '+DBFS_DG'
     SIZE 32G AUTOEXTEND ON NEXT 8G MAXSIZE 300G
     NOLOGGING ONLINE PERMANENT EXTENT MANAGEMENT LOCAL
     AUTOALLOCATE SEGMENT SPACE MANAGEMENT AUTO;
```

- Create a DBFS user account:

```
SQL> create user dbfs identified by dbfs
     quota unlimited on DBFS;
SQL> grant create session, create table, create view,
     create procedure, dbfs_role to dbfs;
```

ORACLE

DBFS is fully functional if it is co-located with the target database. This configuration is preferred when you wish to maintain consistency between your DBFS data store and your other database data. Alternatively, you can configure DBFS in a separate staging database, which can be used by multiple target databases. This configuration is preferred if you want to centralize management of the DBFS data store and de-couple it from other database data.

Use the Database Configuration Assistant (DBCA) to create a database. If you want to access DBFS across multiple Exadata database servers, create the staging database as a RAC database across the desired servers. Default settings are mostly suitable when configuring the staging database. My Oracle Support note 1191144.1 contains the following recommendations:

- Use the **General Purpose or Transaction Processing** template.
- Use Oracle Managed Files and the `DBFS_DG` disk group to store the staging database.
- Deselect **Specify Flash Recovery Area.**
- Set **Memory Size (SGA and PGA)** to 8192 MB.
- Deselect **Use Automatic Memory Management**.
- Use `AL32UTF8` as the database character set.
- Set the `PARALLEL_MAX_SERVERS` initialization parameter to `2`.

Create a dedicated bigfile tablespace to use as the DBFS store. The slide shows an example `CREATE TABLESPACE` command containing the recommended tablespace options. Also, create a DBFS user account and assign privileges as suggested in the slide.

Configuring a DBFS Staging Area

Additional database server operating system configuration:

- Add the user that will mount the DBFS file system to the `fuse` group:

```
# usermod -a -G fuse oracle
```

- As `root`, create `/etc/fuse.conf` containing the entry `user_allow_other`:

```
# echo "user_allow_other" > /etc/fuse.conf
# chmod 644 /etc/fuse.conf
```

- Create a mount point for DBFS with ownership and group permissions set to the user that will mount the DBFS file system:

```
# mkdir /data
# chown oracle:dba /data
```

ORACLE

To configure DBFS, you must also perform the operating system configuration tasks listed in the slide. They should be performed on each database server where DBFS will be mounted. The examples show a typical configuration using the Oracle software owner as the DBFS mount point owner. You can create and use an alternative operating system user if you want to separate DBFS access from the Oracle software owner and its database administration privileges. You can also use the `dcli` utility to replicate the configuration steps on multiple database servers.

Note: The command examples in this lesson focus on database servers running the Linux operating system. A similar procedure can also be used for database servers running the Solaris operating system. See My Oracle Support note 1054431.1 for Solaris-specific commands.

Configuring a DBFS Staging Area

- Creating the DBFS store:

```
$ cd $ORACLE_HOME/rdbms/admin
$ sqlplus dbfs/dbfs
SQL> start dbfs_create_filesystem DBFS mydbfs
```

- Mounting DBFS:

```
$ nohup $ORACLE_HOME/bin/dbfs_client dbfs@<StagingDB>
  -o allow_other,direct_io /data < passwd.txt &
```

  – See My Oracle Support note 1054431.1 for automatic mounting configuration details.

- Using DBFS:
  – Access DBFS through the mount directory.
  – Copy files to DBFS by using methods such as FTP and NFS.

ORACLE

After the staging database is created and the required operating system configuration is completed, you can create the DBFS store. Use the script located at `$ORACLE_HOME/rdbms/admin/dbfs_create_filesystem.sql`. The script must be run by the DBFS database user (created earlier in the configuration process). The script accepts two parameters. In the example in the slide, `DBFS` represents the name of the tablespace that you created to house the DBFS store, and `mydbfs` is the name given to the DBFS store. The name of the DBFS store is used after DBFS is mounted to name the directory that appears under the DBFS mount directory.

To mount DBFS, you can use `dbfs_client` as shown in the example in the slide. Note that the example references `/data` as the DBFS mount directory. Also, note that the DBFS user password is contained in a file called `passwd.txt`. The password file is required only when mounting DBFS and is not required when DBFS is being used.

See My Oracle Support note 1054431.1 for additional details about how to:

- Configure an Oracle Wallet so that the DBFS client can mount a DBFS store without a password.
- Automatically mount DBFS using Oracle Clusterware.

After DBFS is mounted, you can access it through the mount directory (`/data` in the slide example). In the mount directory, you will find a subdirectory named after the DBFS store; the contents of this subdirectory (`/data/mydbfs` in the example) is the contents of the DBFS store.

Note that it is possible to have multiple copies of `dbfs_client` accessing the same shared file system. The sharing and caching semantics are similar to NFS. Like NFS, the default mode caches writes on the client and flushes them after a timeout or when the file being modified is closed. Also like NFS, writes to a file are visible only to clients that open the file after the writer closed the file. This behavior is commonly referred to as close-to-open cache consistency.

For more information about DBFS on Exadata, refer to My Oracle Support note 1054431.1. See also the chapters on DBFS in the *Oracle Database SecureFiles and Large Objects Developer's Guide 11g Release 2 (11.2).*

Configuring the Target Database

- Prerequisites for data file access using external tables:
  - Create an Oracle directory object that references the DBFS staging area directory.
  - Grant the required permissions on the Oracle directory object.
  - Create the required external tables.
- Ensure efficient space management.
  - Use bigfile tablespaces.
  - Use 8 MB initial extents for large segments.
    - Set the `INITIAL` storage parameter to `8 MB`.
    - Set the `CELL_PARTITION_LARGE_EXTENTS` initialization parameter to `TRUE` or `ALWAYS`.
- Use unlimited quotas to bypass quota management.
- Use the parallel clause to set the default degree of parallelism for the target tables.

ORACLE

After the staging area is configured, it can be populated with the required data files. To access the data files using external tables, you must:

- Create an Oracle directory object that references the staging area directory
- Grant the required permissions on the Oracle directory object so that the data files can be referenced by the required target database users
- Create the required external tables that reference the data files

You should also prepare your target database more generally to optimize the load process. You should ensure efficient space management by following the general space management recommendations for use in conjunction with Exadata that are listed in the slide. In addition, you can optimize the performance of your loads if you bypass tablespace quota management and you set an appropriate default degree of parallelism for your target tables.

Loading the Target Database

- The recommended approach uses external tables:
  - Parallel direct path loading for high performance
  - In-flight processing using SQL
    - Transformations using SQL functions
    - Sort data while loading to optimize Exadata storage indexes
  - No need to restage the data
  - Other advanced features
    - Example: Input file preprocessing
- SQL*Loader can also be used:
  - Parallel direct path loading can also be achieved.
  - There is less processing flexibility compared with external tables.
  - Existing SQL*Loader scripts can be easily reused.

ORACLE

The recommended method for loading the data files uses external tables. Loading data using external tables provides the following key benefits:

- You can use direct path loading and parallel processing to achieve high performance.
- You can use the flexibility and power of SQL to efficiently process data while it is being loaded. For example, you can use standard or user-defined SQL functions to transform the data during loading and you can sort the data while you load it to optimize the storage indexes that are automatically maintained by Exadata Storage Server.
- Because you can process data while it is being loaded, it is unlikely that you will need to restage the data inside the target database.
- External tables provide some useful advanced features. A primary example is the ability to preprocess a data file using a user-defined program, which provides additional flexibility and power to process the data using routines outside the database.

SQL*Loader can also be used to load data files. It can deliver comparable performance, but does not provide the same level of processing flexibility when compared with external tables. SQL*Loader remains a good choice when customers already have SQL*Loader–based scripts that they want to reuse in an Exadata environment.

Loading the Target Database

- Parallel loading using external tables:
  - A `CREATE TABLE ... AS SELECT` statement automatically uses the default degree of parallelism.
  - An `INSERT ... SELECT` statement needs parallel DML to be enabled:

```
SQL> alter session enable parallel dml;
```

- Direct path loading using external tables
  - A `CREATE TABLE ... AS SELECT` statement automatically uses direct path loading.
  - An `INSERT ... SELECT` statement needs an `APPEND` hint to enable direct path loading:

```
SQL> insert /*+ APPEND */ into my_table
     select * from my_external_table;
```

ORACLE

Bulk data loading using external tables is achieved by using a `CREATE TABLE ... AS SELECT` or an `INSERT ... SELECT` statement.

To use parallel processing, it is recommended that you set the default degree of parallelism for the external tables used in the load and the target table being created or inserted into. Alternatively, you can set parallelism within the statement using a `PARALLEL` hint.

After parallelism is configured, a `CREATE TABLE ... AS SELECT` statement will automatically use the defined degree of parallelism. An `INSERT ... SELECT` statement requires parallel DML to be enabled before it can execute in parallel.

To use direct path loading from external tables, you must use an `APPEND` hint for an `INSERT ... SELECT` statement. `CREATE TABLE ... AS SELECT` statements automatically use direct path loading.

Remember that the bulk data loading approach outlined in this lesson can be used in conjunction with other techniques, such as partition exchange loading.

Quiz

To facilitate parallel loading from typical flat files, external table definitions must reference numerous smaller data files rather than one file containing all the data.

a. True
b. False

ORACLE

Answer: b

Typically, Oracle automatically divides typical flat files into 10 MB granules to facilitate parallel processing. Oracle is unable to use this approach with compressed files or data read from a pipe or a tape device.

Quiz

Although DBFS is fully functional if it is collocated with the target database, it is recommended to configure DBFS in a separate staging database.

a. True

b. False

Answer: a

Summary

In this lesson, you should have learned how to:

- Describe alternative architectures for bulk data loading into Exadata
- Configure the Database File System (DBFS) feature for staging input data files
- Use external tables to perform high-performance data loads

ORACLE

Additional Resources

- My Oracle Support notes
  - Configuring a Database for DBFS on Oracle Database Machine https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1191144.1
  - Configuring DBFS on Oracle Database Machine https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1054431.1

Practice 12 Overview: Bulk Data Loading Using Oracle DBFS

In this practice, you will perform a bulk data load on Exadata. You will configure a database file system (DBFS) and use it to stage a CSV-formatted file. You will then use the external table feature of Oracle Database to reference the CSV file. Finally, you will use a `CREATE TABLE AS SELECT` statement to copy the CSV file data into a table in your database.

13

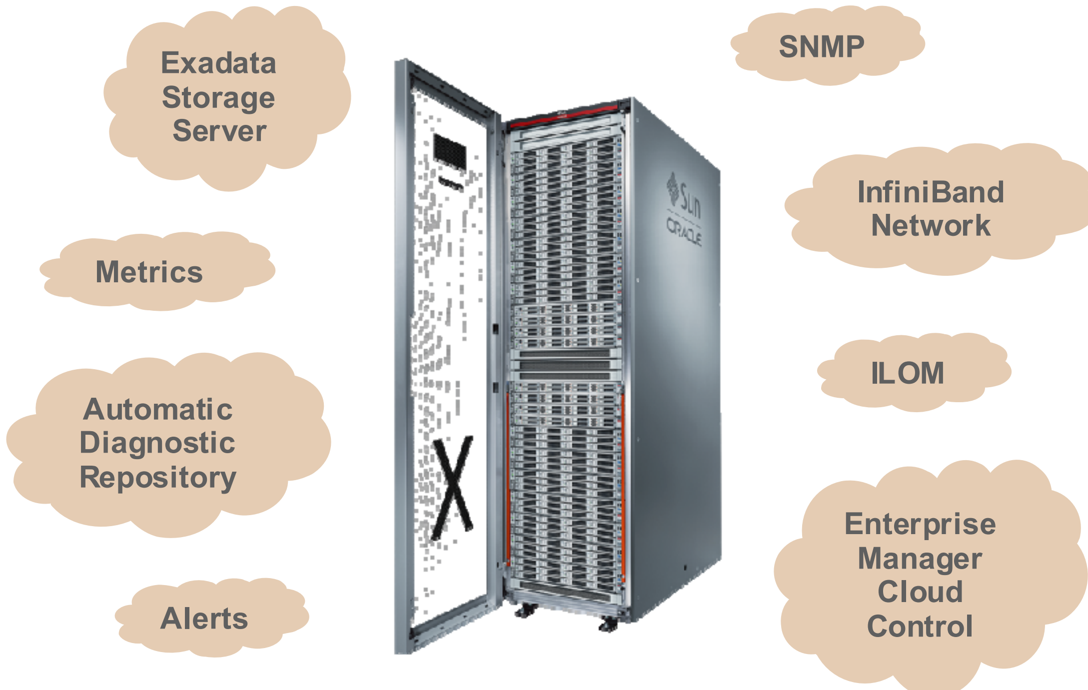
Exadata Database Machine Platform Monitoring: Introduction

ORACLE

Objectives

After completing this lesson, you should be able to describe the key monitoring infrastructure technologies associated with Exadata Database Machine.

Monitoring Technologies and Standards



A variety of monitoring technologies are embedded in the various Exadata Database Machine components. Some are based on widely adopted standards, whereas others are specific to Database Machine. This lesson introduces the key monitoring technologies and standards used in conjunction with Database Machine. Most of the information presented in this lesson is familiar to most students and is presented to provide a consistent foundation for the following lessons.

Simple Network Management Protocol (SNMP)

- What is it?
  - A standard protocol for managing devices on a network
- What does it do?
  - Primarily used to propagate information about a device to a monitoring console somewhere on the network
    - Report alerts for hardware or software issues
  - Can be used to set device configuration parameters
- Where is it found?
  - Throughout Database Machine, including database servers, storage servers, InfiniBand switches, Ethernet switch, and Power Distribution Units (PDUs)

ORACLE

Simple Network Management Protocol (SNMP) is an Internet-standard protocol for managing devices on IP networks. A wide array of devices, including many servers, network switches, routers, printers, and workstations, support SNMP. SNMP is used mostly in network management systems to monitor network-attached devices for conditions that warrant administrative attention.

Typically, SNMP is used to propagate information about a device to a monitoring or management console located somewhere in the network. To do this, managed devices run an SNMP agent. SNMP agents expose management information about the managed systems as variables. The protocol also permits active management tasks, such as modifying and applying a new configuration through remote modification of these variables. SNMP agents can be separate processes or embedded into other software or hardware modules.

SNMP defines a variety of protocol data units (messages) that can be exchanged between agents and managers. One of the most important messages is an asynchronous notification from agent to manager known as a trap. An SNMP trap is often used by devices to report alert conditions.

SNMP is used extensively throughout Exadata Database Machine so that the various components of Database Machine can report monitoring information and alerts to network management systems, such as Enterprise Manager Cloud Control.

Intelligent Platform Management Interface (IPMI)

- What is it?
  - An open, industry-standard interface for server management
- What does it do?
  - Primarily used to perform server configuration and management independently of the server operating system
- Where is it found?
  - Exadata Database Machine database servers and Exadata Storage Servers

ORACLE

Intelligent Platform Management Interface (IPMI) is an open, industry-standard interface for the management of server systems over several different types of networks.

Although it seems that SNMP and IPMI perform the same function, there are two main differences:

- IPMI is focused on server management. IPMI functionality includes field-replaceable unit (FRU) inventory reporting, logging of system events, and system recovery (including system resets, power on, and power off).
- IPMI is associated with an architecture for administrators to remotely manage a system in the absence of an operating system or other system management software. The monitored system may be powered off, but the baseboard management controller (BMC) must be connected to a power source and the monitoring medium, typically a local area network connection. The BMC is a specialized microcontroller embedded in the server.

IPMI prescribes only the structure and format of the interfaces as a standard. The detailed implementations may vary.

Inside Database Machine, IPMI support is built into Integrated Lights Out Manager (ILOM) on each database server and Exadata Storage Server.

Integrated Lights Out Manager (ILOM)

- What is it?
  – An integrated service processor hardware and software
- What does it do?
  – Provides out-of-band server monitoring and management to:
    – Remotely control the power state of a server
    – View the status of sensors and indicators on the system
    – Provide a remote server console
  – Generates alerts for hardware errors and faults as they occur
- Where is it found?
  – Exadata Database Machine database servers and Exadata Storage Servers

ORACLE

Integrated Lights Out Manager (ILOM) provides advanced service processor hardware (baseboard management controller) and software that you can use to manage and monitor Database Machine servers. ILOM's dedicated hardware and software is preinstalled on the database servers and Exadata Storage Servers inside Database Machine. With ILOM, you can actively manage and monitor the server independently of the operating system state (out-of-band). You can:

- Learn about hardware errors and faults as they occur
- Remotely control the power state of your server
- View and use the graphical and non-graphical consoles for the host
- View the current status of sensors and indicators on the system
- Determine the hardware configuration of your system
- Receive alerts about important system events through notification methods, such as SNMP traps and email alerts

ILOM automatically initializes as soon as power is applied to the server. It provides a full-featured, browser-based web interface and has an equivalent command-line interface (CLI). There is also an industry-standard SNMP interface and IPMI support. (ILOM is compliant with IPMI v1.5 and v2.0.)

Exadata Storage Server: Metrics, Thresholds, and Alerts

- What are they?
  - Metrics, thresholds, and alerts provide the foundation for monitoring Exadata Storage Server.
- What do they do?
  - Metrics provide a measure relating to some aspect of storage server status or performance.
  - Thresholds are metric levels, which if crossed automatically, generate an alert notification.
  - Alerts are automatically generated notifications of system events.
- Where are they found?
  - Exadata Storage Server

ORACLE

Exadata Storage Server contains a system of metrics, thresholds, and alerts, which provide a foundation for storage server monitoring.

Metrics provide a measure relating to some aspect of storage server status or performance. Exadata Storage Server includes more than 150 metrics, many of which are associated with numerous separate measurements. For example, the current temperature of the server is a single-value metric. In contrast, the cumulative number of requests to read large blocks from a grid disk has a separate measurement for each grid disk. So, on a typical Exadata cell, this one metric can be associated with more than 20 separate observations.

Thresholds are definitions that enable administrators to define metric levels, which if crossed automatically generate an alert notification. Thresholds can contain two alert levels: warning and critical. For example, a threshold could be defined to generate a warning alert when cell memory utilization reaches 80 percent, and a critical alert when it reaches 90 percent.

Exadata Storage Server generates alerts for various system conditions. For example, an alert is generated if the cell server software terminates unexpectedly or if a sensor detects a potential imminent disk failure. Additional alerts can be generated for user-defined events by using thresholds.

Note that Enterprise Manager also contains a separate system of metrics, thresholds, and alerts. Other software, such as Oracle Database, contains its own metrics and alerts.

Automatic Diagnostic Repository (ADR)

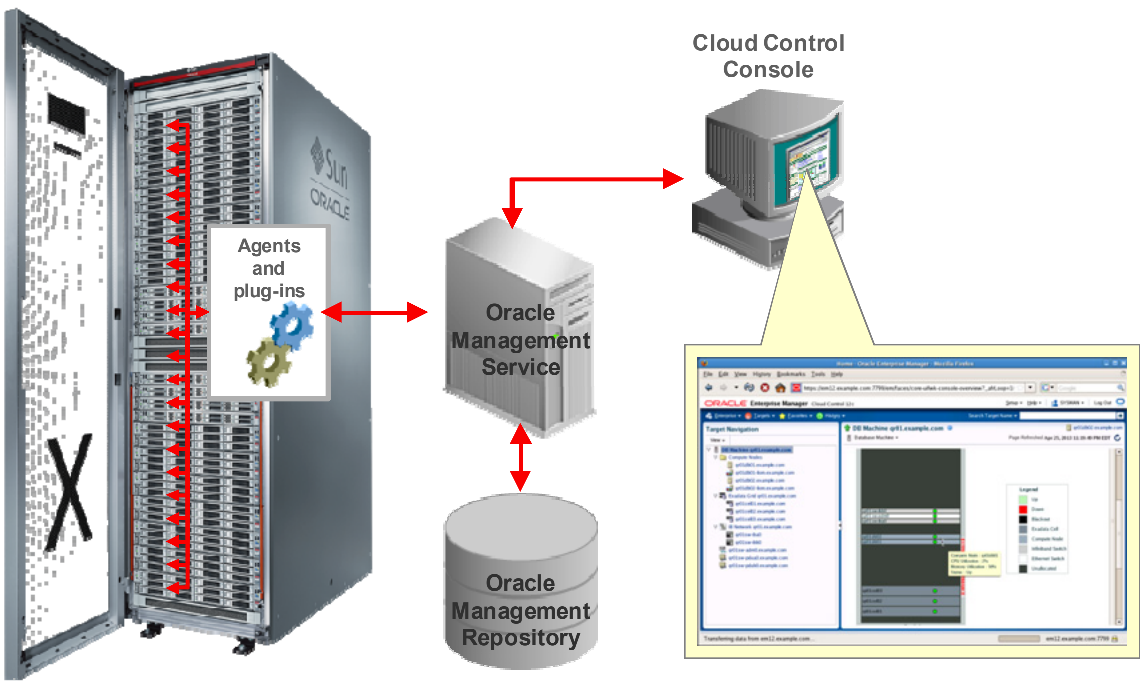
- What is it?
  - A file-based repository for diagnostic data such as traces, dumps, and logs
- What does it do?
  - Provides a consistent organization for diagnostic data, which enables administrators and Oracle Support to correlate and analyze diagnostic data more easily and effectively
- Where is it found?
  - Exadata Database Machine database servers and Exadata Storage Servers

The Automatic Diagnostic Repository (ADR) is a file-based repository for diagnostic data, such as traces, dumps, the alert log, health monitor reports, and more. It has a unified directory structure across multiple products and multiple instances. Beginning with Release 11*g*, Oracle Database, Automatic Storage Management (ASM), Oracle Net Listener, and other Oracle products store diagnostic data in the ADR. Exadata Storage Server also uses the ADR structure to organize its diagnostic data.

Each instance of each product stores diagnostic data under its own home directory within the ADR. For example, in an Oracle Real Application Clusters environment with shared storage and Oracle ASM, each database instance and each ASM instance have separate ADR home directories.

ADR's unified directory structure, consistent diagnostic data formats across products and instances, and a unified set of tools enable customers and Oracle Support to efficiently correlate and analyze diagnostic data across multiple instances.

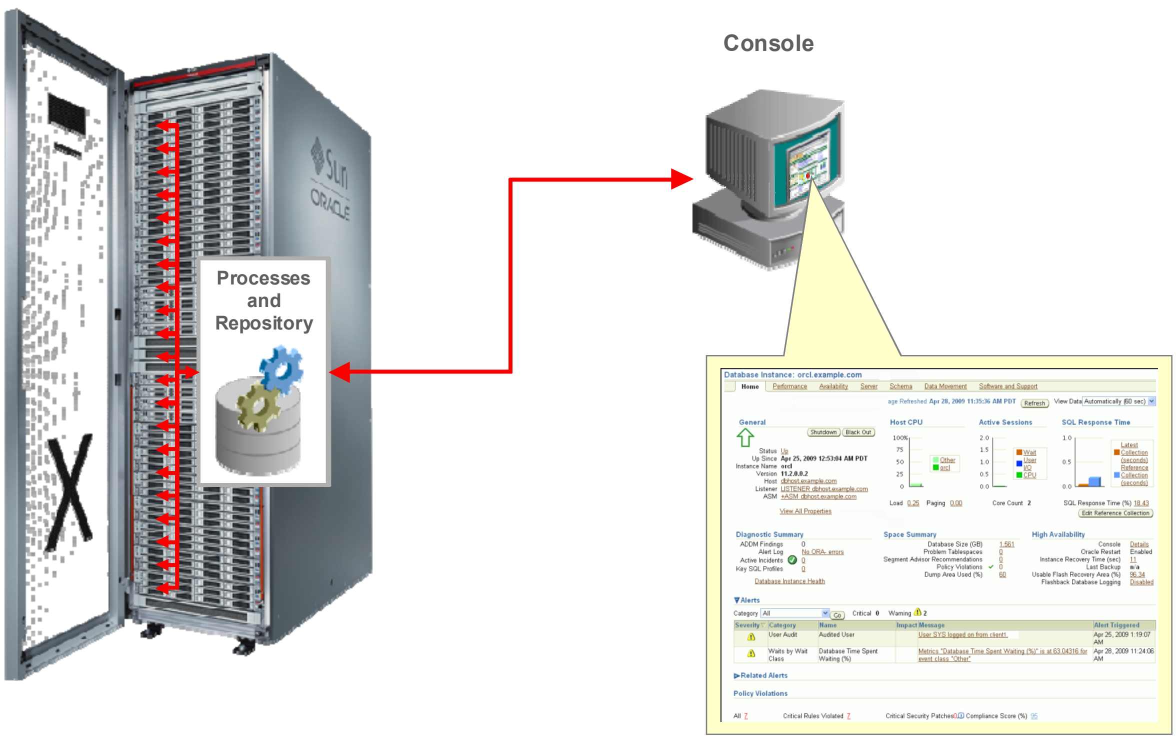
Enterprise Manager Cloud Control



Enterprise Manager Cloud Control is a system management software platform that delivers centralized monitoring, administration, and lifecycle management functionality for information technology infrastructure, including systems running Oracle and non-Oracle technologies.

When used in conjunction with Exadata Database Machine, Enterprise Manager Cloud Control should be configured with the Oracle Exadata plug-in to enable monitoring of the various Database Machine components. Additionally, there are best practice recommendations regarding how to configure Database Machine targets inside Cloud Control. For information about configuring Cloud Control for monitoring Exadata Database Machine, refer to the lesson titled “Configuring Enterprise Manager Cloud Control to Monitor Exadata Database Machine.”

Enterprise Manager Database Control and Enterprise Manager Database Express



Enterprise Manager Database Control is a database administration environment built in to Oracle Database 11*g*. Enterprise Manager Database Express (EM Express) is the next-generation management tool for administration and management of individual databases that comes with Oracle Database 12*c* Release 1.

Both environments provide a user-friendly interface to perform core administration and monitoring tasks on the database. However, neither environment provides the breadth of monitoring capabilities offered by Cloud Control. For example, neither environment provides any InfiniBand network monitoring capabilities. For this reason, Cloud Control is the recommended monitoring environment for Exadata Database Machine.

Database Control and EM Express are not considered further during this course.

Quiz

Which of the following monitoring infrastructure technologies can be used in conjunction with Exadata Storage Servers?

a. Simple Network Management Protocol (SNMP)
b. Integrated Lights Out Manager (ILOM)
c. Exadata metrics, thresholds, and alerts
d. Automatic Diagnostic Repository (ADR)
e. Enterprise Manager Cloud Control

Answer: a, b, c, d, e

Summary

In this lesson, you should have learned how to describe the key monitoring infrastructure technologies associated with Exadata Database Machine.

14

Configuring Enterprise Manager Cloud Control to Monitor Exadata Database Machine

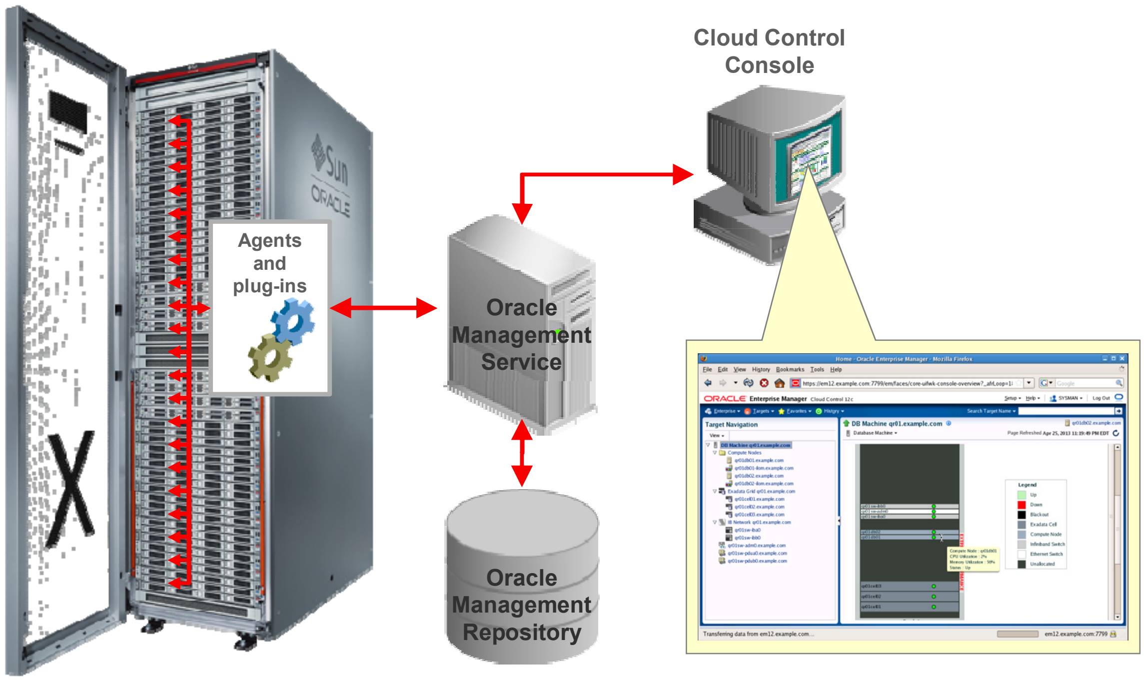
Objectives

After completing this lesson, you should be able to:

- Describe Enterprise Manager Cloud Control architecture as it specifically applies to Exadata
- Explain how to configure Enterprise Manager Cloud Control to monitor Exadata
- Explain the guided discovery process for Exadata
- Describe how to configure a dashboard for Exadata

ORACLE

Enterprise Manager Cloud Control Architecture: Overview



Enterprise Manager Cloud Control contains four major elements:

- **Oracle Management Service (OMS):** The processing heart of the system
- **Oracle Management Repository:** The persistent store of monitoring, management, and configuration information
- **Cloud Control Console:** A web-based management interface
- **Agents and plug-ins:** Deployed to managed targets to collect information, which is processed by the OMS

This fundamental architecture does not change when Cloud Control is used to monitor Exadata.

Enterprise Manager Cloud Control: Supported Exadata Configurations

- Supported Exadata configurations:
  - Multi-rack Exadata systems including:
    - All V2, X2, X3, X4, X5 and X6 models
    - Exadata Storage Expansion Racks
    - Compute Node Expansion Rack
  - Partitioned Exadata systems:
    - Logically splitting an Exadata rack into multiple separate environments
- Unsupported configurations:
  - Exadata V1

ORACLE

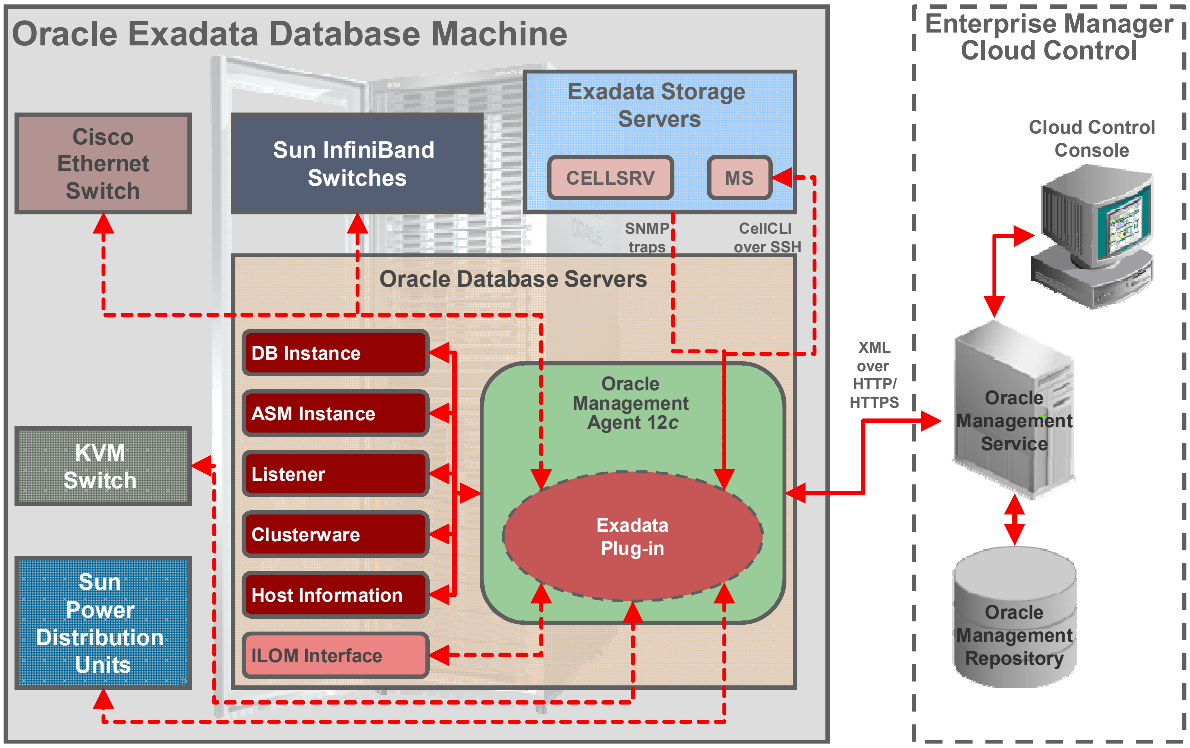
The slide lists the Exadata configurations that are supported by Enterprise Manager Cloud Control.

Note that Cloud Control provides support for partitioned Exadata environments only when:

- The partitions are all defined and installed at the same time, not through multiple invocations of the Exadata deployment tool.
- Cells and compute nodes are not shared between partitions
- Multiple partitions are connected through the same InfiniBand network
- Compute nodes in the same partition share the same Cluster

Enterprise Manager Cloud Control does not provide support for Exadata V1.

Cloud Control Monitoring Architecture for Exadata



When Enterprise Manager Cloud Control is used to monitor Exadata, an Oracle Management Agent and Oracle Exadata plug-in should be deployed to every database server.

The agent is used to monitor and maintain Oracle software targets on the database server, including database instances, ASM instances, listeners, clusterware resources, and other host information. The Oracle Exadata plug-in extends Cloud Control to enable monitoring of other Exadata components, such as the Exadata Storage Servers, InfiniBand switches, Ethernet switch, Keyboard, Video and Mouse (KVM) switch (installed on some earlier models), and Power Distribution Units (PDUs).

The Exadata plug-in connects to its monitoring targets using different methods. For example, the plug-in gathers monitoring data from the Exadata Storage Servers by making CellCLI calls over SSH. However, Exadata cell alerts are delivered directly to the agent by using SNMP traps.

Configuring Cloud Control to Monitor Exadata

Prerequisites: Install and configure Exadata and Enterprise Manager Cloud Control.

1. Perform prediscovery configuration and verification.
2. Deploy Oracle Management Agent to all database servers.
3. Discover Exadata.
4. Perform postdiscovery configuration and verification.
5. Discover additional targets.
6. Configure an Exadata dashboard.

The slide outlines the recommended steps to configure Enterprise Manager Cloud Control for monitoring Exadata. The assumed starting point is after the initial installation and configuration of Exadata and Enterprise Manager Cloud Control with the Exadata plug-in.

The remainder of this lesson describes each step in greater detail. You will also configure Enterprise Manager Cloud Control for monitoring Exadata in the practices that accompany this lesson.

Pre-Discovery Configuration and Verification

- Create dedicated ILOM user IDs on the database servers.
- Verify software versions:
  - Exadata Storage Server
  - ILOM `ipmitool`
  - InfiniBand Switch firmware
  - PDU firmware
  - KVM application (if installed)
- Verify that Oracle Grid Infrastructure and Oracle Database are running.
- Verify host name resolution:
  - OMS to database servers
  - Database servers to other components
- Verify access to network ports and services, and modify firewalls.
- Configure Enterprise Manager roles and users.

ORACLE

The slide outlines the prediscovery configuration and verification tasks that should be performed before implementing Enterprise Manager in conjunction with Exadata. Detailed instructions for each task are documented in chapter 2 of the *Oracle Enterprise Manager Exadata Management Getting Started Guide Release 12.1*.

Enterprise Manager requires ILOM user credentials to communicate with an ILOM server processor. The ILOM administrator (`root`) user account can be used. However, it is recommended that you create a dedicated ILOM user account that is associated with the operator role (`cro`).

The software on various Exadata components must be compatible with Enterprise Manager and the Exadata plug-in. *Oracle Enterprise Manager Exadata Management Getting Started Guide Release 12.1* contains details about the required versions and verification instructions. Note that the software version requirements may change over time. Therefore, always consult the latest available documentation.

Before discovery, it is recommended that Oracle Grid Infrastructure software is up and running, and that your Oracle databases are started.

The Enterprise Manager OMS server must be able to resolve the host name for each of the database servers. Also, each database server should be verified to ensure that it is able to resolve the host names of the ILOM servers, PDUs, storage cells, InfiniBand switches, Cisco Ethernet switch, and the KVM switch (if present).

Enterprise Manager relies on various network services to monitor and manage Exadata. You will need to adjust your firewall settings to allow for the following:

- Enterprise Manager uses `ping` to establish the basic availability and status of the Exadata components. Specifically, the OMS server must be able to ping the Exadata database servers, and the Exadata database servers hosting the Enterprise Manager Agents must be able to ping the other Exadata components (ILOM servers, PDUs, storage cells, InfiniBand switches, Cisco switch, and KVM switch).
- The database listener ports must be open to the OMS server. Note that Exadata databases use SCAN listeners. So, ports need to be open for each database server client network interface, each database server VIP, and the SCAN VIPs.
- The Enterprise Manager Agents require access to the upload service on the OMS server, which is normally configured on port 4889 for HTTP uploads and 4900 for HTTPS.
- The OMS server must be able to connect to the Enterprise Manager Agent HTTP/HTTPS port on each compute node. The agent port defaults to 3872.
- The Enterprise Manager Agents require SSH access to the Exadata components that they monitor (ILOM servers, PDUs, storage cells, InfiniBand switches, Cisco switch, and KVM switch). Note that if your Exadata Storage Servers are configured to lock out access by using SSH, you must ensure that you use a version of Enterprise Manager that supports this configuration.
- All Exadata components need to be able to send SNMP traps (using UDP) to the agents running on the database servers.

Enterprise Manager contains a system of roles and privileges that enables controlled access to different monitoring and administration functions. It is recommended to create specific roles and users to monitor and administer Exadata, rather than using `SYSMAN`.

Deploying the Oracle Management Agent

- Where to deploy:
  - All Exadata database servers
- Recommended version:
  - 12*c* Release 1 version 12.1.0.2 or later
- How to deploy:
  - During the initial configuration of Exadata
  - Push Method:
    - Interactively using the Cloud Control Agent Installation Wizard
    - Silent mode also available
  - Automation Kit:
    - Deploys all agents using a single command
    - See My Oracle Support bulletin 1440951.1

To facilitate management of Exadata using Enterprise Manager Cloud Control, you must deploy the Oracle Management Agent. It is recommended to deploy Oracle Management Agent 12*c* Release 1 version 12.1.0.2 (or later) on all the Exadata database servers.

Agent deployment can be performed during the initial configuration of Exadata. This option was discussed earlier in the lesson titled “Exadata Database Machine Initial Configuration.”

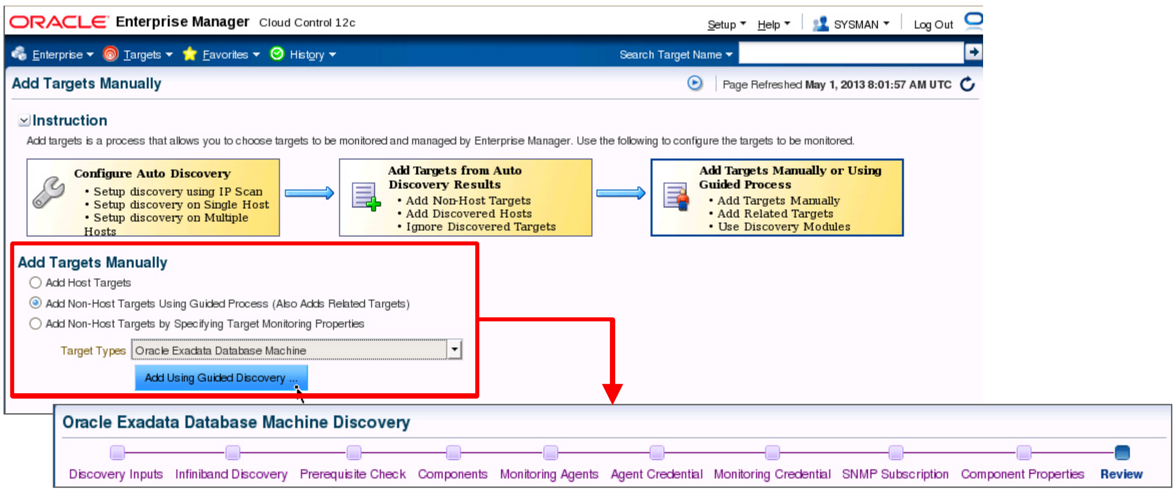
Otherwise, assuming that Enterprise Manager Cloud Control is already running and there is network connectivity between it and the Database Machine, the easiest way to deploy the agent is to use the Cloud Control Agent Installation Wizard. A silent mode option is also provided, which requires you to use a response file to provide all of the required deployment parameters.

Alternatively, an automation kit is provided through Oracle Support that enables agent deployment across your Exadata environment using a single command. The automation kit is available through My Oracle Support bulletin 1440951.1.

The options for agent deployment are discussed in detail in chapter 2 of the *Oracle Enterprise Manager Exadata Management Getting Started Guide Release 12.1*.

Discovering Exadata

- Verify the prerequisites:
  - Run the discovery precheck script.
  - Perform additional manual verification tasks.
- Perform Exadata guided discovery.



Before embarking on the Exadata discovery process, there are a few checks that you should perform to ensure a smooth discovery. These checks are documented in chapter three of the *Oracle Enterprise Manager Exadata Management Getting Started Guide Release 12.1*.

After the prerequisites are met, you can use Enterprise Manager to perform discovery. To start, select the Setup > Add Targets > Add Targets Manually menu command.

Then on the Add Targets Manually page:

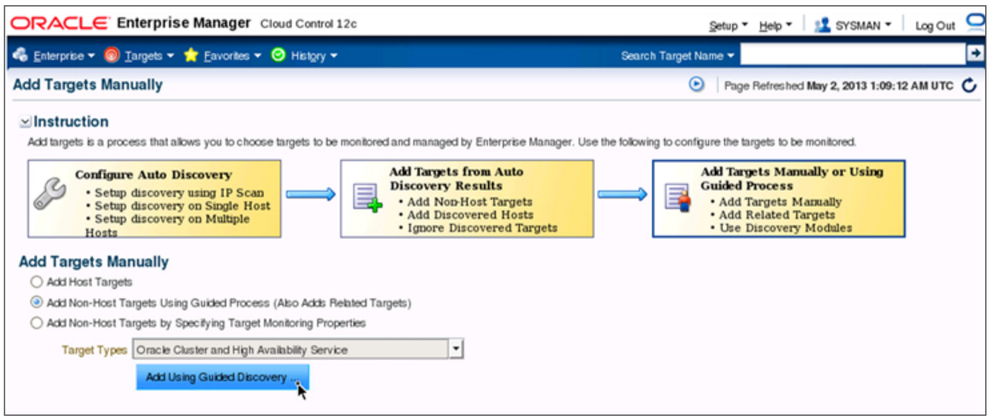
- Select the Add Non-Host Targets Using Guided Process (Also Adds Related Targets) option
- From the Target Types list, select Oracle Exadata Database Machine
- Click Add Using Guided Discovery to start the multistep wizard that will guide you through the discovery process

The slide shows an example of the Add Targets Manually page, along with the steps in the Exadata Discovery Wizard.

The steps in the Exadata Discovery Wizard, along with the inputs required by each step, are described in chapter three of the *Oracle Enterprise Manager Exadata Management Getting Started Guide Release 12.1*. You will also perform the discovery process in the practice associated with this lesson.

Discovering Additional Targets

- Perform additional guided discovery for:
  - Oracle Cluster and High Availability Service
  - Oracle Database, Listener, and ASM
- Procedures are essentially the same as for non-Exadata systems:



Following the Exadata guided discovery process, Enterprise Manager is able to monitor and manage the Exadata system components, including the database servers, Exadata storage servers, InfiniBand switches, Cisco Ethernet switch, Oracle ILOM servers, PDUs, and KVM switch (if present).

Additional configuration is required to complete the discovery of the various Oracle Grid Infrastructure and Oracle Database software components that run on Exadata. The quickest and easiest way to discover the software components is to use the guided discovery processes for:

- Oracle Cluster and High Availability Service
- Oracle Database, Listener, and Automatic Storage Management

You should always discover the cluster first to enable proper discovery of Oracle Real Application Clusters (RAC) databases.

The procedures for cluster discovery and database discovery are essentially the same, regardless of whether the components reside on Exadata or not. Therefore, Oracle database administrators who are familiar with Enterprise Manager should already be familiar with these tasks.

Post-Discovery Configuration and Verification

- Configure agent hosts to forward SNMP notifications.
- Verify SNMP settings for storage cells and InfiniBand switches.
- Configure and verify SNMP settings for:
  - Oracle ILOM Server targets
  - Cisco Ethernet Switch targets
  - Power Distribution Unit (PDU) targets
  - KVM switch targets

ORACLE

The slide outlines the configuration and verification tasks that should be performed after the Exadata guided discovery process. Detailed instructions for each task are documented in chapter 4 of the *Oracle Enterprise Manager Exadata Management Getting Started Guide Release 12.1*.

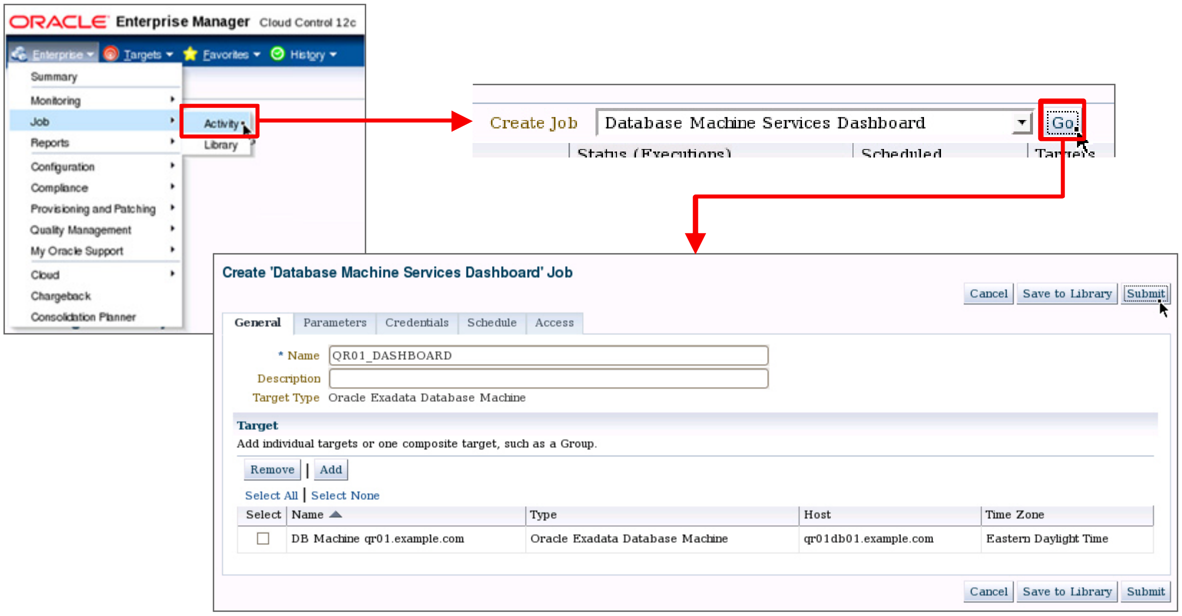
Some Exadata components, such as the Cisco Ethernet switch, have no documented way to send SNMP traps to ports other than UDP 162. However, the Oracle Management Agent runs as the `oracle` user and cannot listen on UDP ports under 1024. Because of these restrictions, a trap forwarder must be configured so that the SNMP traps sent to port 162 are forwarded to Enterprise Manager through the Oracle Management Agent.

During the guided discovery process, you have the option to allow Enterprise Manager to configure SNMP subscriptions for the Exadata Storage Servers and the InfiniBand switches. If you did not select these options, you will need to manually configure the SNMP subscriptions to enable these components to send alert information to Enterprise Manager. If you did select the automatic SNMP subscription option, you should check the cells and InfiniBand switches to confirm the configured settings. You should also confirm that any trap settings apart from those relating to Enterprise Manager continue to function as desired.

Finally, you must also manually configure the SNMP subscription settings for the Oracle ILOM servers, Cisco Ethernet switch, PDUs, and KVM switch (if present).

Configuring an Exadata Dashboard

Configure an Exadata dashboard to monitor Exadata hardware and software components:



ORACLE

The Database Machine Services Dashboard provides a one-stop overview that enables you to quickly and easily monitor the key availability and performance metrics for your Exadata system and the related software services on one page. Although you can monitor Exadata without a dashboard, it is recommended that you create a Database Machine Services Dashboard.

To begin the dashboard creation process, select the Enterprise > Jobs > Activity menu command. Then in the Create Job list, select the option to create a Database Machine Services Dashboard and click Go. On the resulting page, specify the dashboard job name and specify the Database Machine associated with the dashboard. Finally, submit the job.

After the job completes, the dashboard is ready for use. To access the dashboard, select the Enterprise > Reports > Information Publisher Reports menu command. On the resulting page, locate the dashboard and click the link associated with its name. Database Machine Services Dashboards are named in the following format:

```
<Database Machine Name>_DASHBOARD_REPORT
```

Quiz

The Oracle Management Agent should be deployed to every Exadata database server because:

a. An agent is required to monitor the cluster and database software components on that database server
b. Multiple agents communicate with each other to cooperatively monitor the other Exadata system components
c. Multiple agents enhance the availability and performance of Exadata monitoring

Answer: a, c

Each database server requires an agent to monitor the cluster and database software components on that server. However, each agent is independent of all other agents, and agents do not communicate with each other to monitor and manage other Exadata components. Rather, each agent is directed to monitor a subset of Exadata components, and the responsibility for monitoring different components can be passed to another agent if the primary monitoring agent is unavailable for any reason. Thus, multiple agents enhance the availability and performance of Exadata Monitoring.

Quiz

The Exadata discovery process configures Enterprise Manager to monitor all Exadata hardware and software components.

a. True
b. False

Answer: b

The Exadata discovery process configures Enterprise Manager to monitor Exadata system components. Separate discovery processes are required to configure Oracle Grid Infrastructure and Oracle Database software components within Enterprise Manager.

Quiz

The discovery processes for Oracle Grid Infrastructure and Oracle Database software components are essentially the same regardless of whether Exadata is involved or not.

a. True

b. False

Answer: a

The procedures for cluster discovery and database discovery are essentially the same regardless of whether the components reside on Exadata or not. Therefore, Oracle database administrators who are familiar with Enterprise Manager should already be familiar with these tasks.

Summary

In this lesson, you should have learned to describe:

- Enterprise Manager Cloud Control architecture as it specifically applies to Exadata
- How to configure Enterprise Manager to monitor Exadata
- The guided discovery process for Exadata
- How to configure a dashboard for Exadata

Additional Resources

- My Oracle Support notes
  - Oracle Database Machine Monitoring Best Practices
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1110675.1

Practice 14 Overview: Exadata Monitoring Configuration

In these practices, you will configure Enterprise Manager Cloud Control to monitor Exadata. You will also perform a selection of post-discovery configuration and verification tasks.

15

Monitoring Exadata Storage Servers

ORACLE

Objectives

After completing this lesson, you should be able to:

- Describe Exadata Storage Server metrics, alerts, and active requests
- Identify the recommended focus areas for Exadata Storage Server monitoring
- Describe how to monitor the recommended Exadata Storage Server focus areas

ORACLE

Lesson Overview

- Exadata Storage Server Monitoring Architecture
- Monitoring Exadata Storage Server with Enterprise Manager: Overview
- Monitoring Hardware Failure and Sensor State
- Monitoring Exadata Storage Server Availability
- Checking for Undelivered Alerts
- Checking for Disk I/O Errors
- Checking for Network Errors
- Monitoring File System Free Space
- Comparing Metrics Across Multiple Storage Servers
- Monitoring Metrics Within a Storage Server
- Third-Party Monitoring Tools

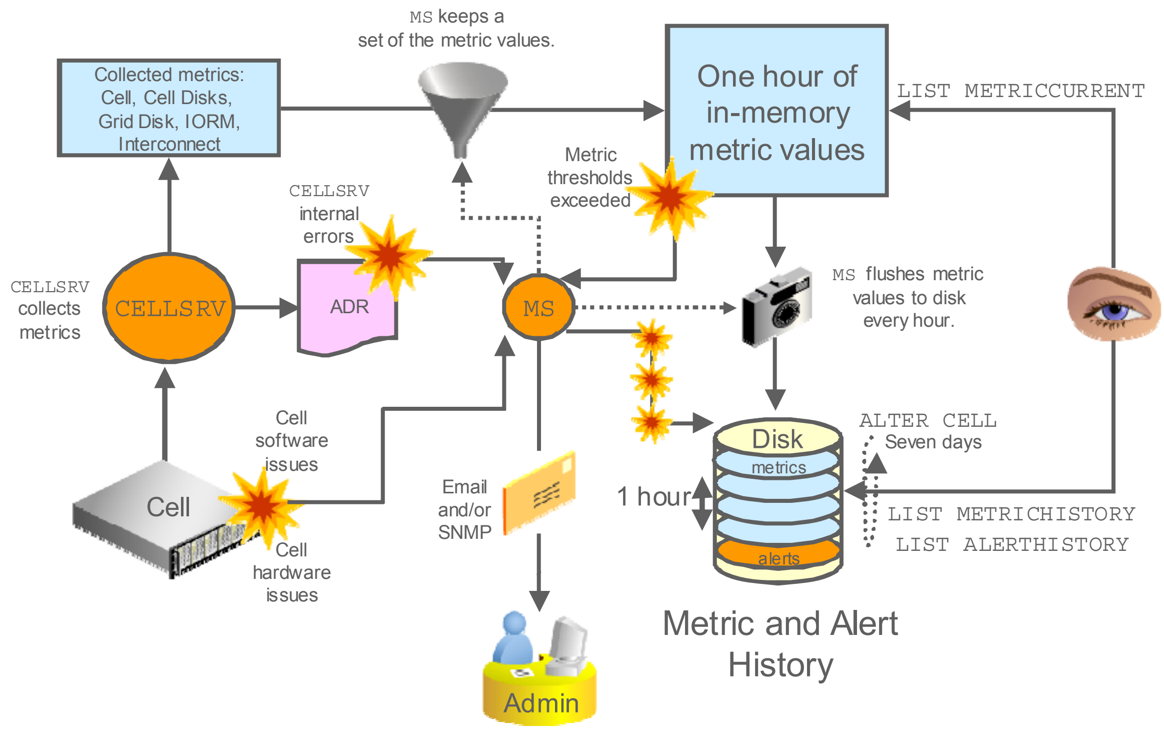
What to Monitor

ORACLE

This lesson focuses on the best practice recommendations for Exadata Storage Server about what should be monitored from a general administrative perspective, and how it should be monitored.

The list in the slide summarizes the topics contained in the lesson. The shaded area highlights the recommended areas to monitor.

Exadata Storage Server Metrics and Alerts Architecture



The diagram in the slide illustrates the architecture for Exadata Storage Server metrics and alerts. It shows how CELLSRV periodically records important runtime properties, called metrics, for cell components, such as CPUs, cell disks, grid disks, flash cache, and I/O Resource Management (IORM) statistics. These metrics are recorded in memory. Based on its own metric collection schedule, the Management Server (MS) gets the set of metric data accumulated by CELLSRV. MS keeps a subset of the metric values in memory, and writes a history to an internal disk-based repository every hour. This process is conceptually similar to database AWR snapshots.

The retention period for metric and alert history entries is specified by the `metricHistoryDays` cell attribute. You can modify this setting with the CellCLI `ALTER CELL` command. By default, it is seven days. You can view the metric value history by using the CellCLI `LIST METRICHISTORY` command, and you can view the current metric values by using the `LIST METRICCURRENT` command.

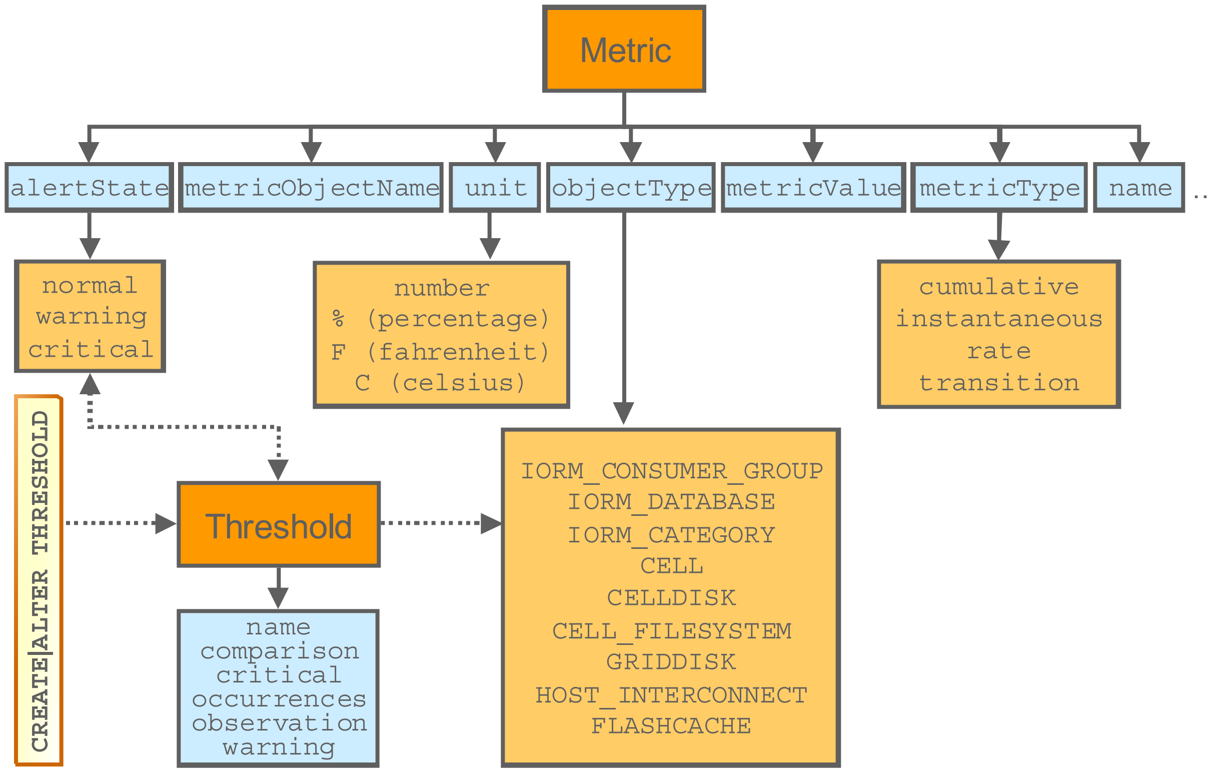
In addition to collecting metrics, Exadata Storage Server can generate alerts. Alerts represent events of importance occurring within a cell, often indicating that an Exadata cell function is compromised.

MS generates an alert when it discovers a:

- Cell hardware issue
- Cell software or configuration issue
- CELLSRV internal error
- Metric that has exceeded a threshold defined in the cell

You can view previously generated alerts by using the `LIST ALERTHISTORY` command. In addition, you can configure the cell to automatically send an email and/or SNMP message to a designated set of administrators.

Monitoring Exadata Storage Server with Metrics



Metrics are recorded observations of important runtime properties or internal instrumentation values of the storage cell and its components, such as cell disks or grid disks. Metrics are a series of measurements that are computed and retained in memory for a period of time, and stored on a disk for a more permanent history.

The graphic in the slide lists some of the important metric attributes. Each metric:

- Has a name and description
- Is associated with a `metricObjectName`, which is the name of the object being measured, such as a specific cell disk, grid disk, or consumer group
- Belongs to a group that is defined by its `objectType` attribute. The possible groups are shown in the slide.
- Has a `metricType`, which is an indicator of how the statistic was created or defined. Possible values and their meanings are:
  - **`cumulative`**: Cumulative statistics since the metric was created
  - **`instantaneous`**: Value at the time that the metric is collected
  - **`rate`**: Rates computed by averaging statistics over observation periods
  - **`transition`**: Collected when the metric value changes, and typically captures important transitions in hardware status
- Has a measurement `unit`. Possible units are shown in the slide.

Understanding the composition of the metric name provides a good insight into the meaning of the metric. The value of the name attribute is a composite of abbreviations. The attribute value starts with an abbreviation of the object type on which the metric is defined:

- `CL_` (cell)
- `CD_` (cell disk)
- `GD_` (grid disk)
- `FC_` (flash cache)
- `DB_` (database)
- `CG_` (consumer group)
- `CT_` (category)
- `N_` (interconnect network)

After the abbreviation of the object type, many metric names conclude with an abbreviation that relates to the description of the metric. For example, `CL_FANS` is the instantaneous number of working fans on the cell.

I/O-related metric name attributes continue with one of the following combinations to identify the operation:

- `IO_RQ` (number of requests)
- `IO_BY` (number of MB)
- `IO_TM` (I/O latency)
- `IO_WT` (I/O wait time)

Next in the name could be `_R` for read or `_W` for write. Following that, there might be `_SM` or `_LG` to identify small or large I/Os, respectively. At the end of the name, there could be `_SEC` to signify per second or `_RQ` to signify per request as in the following examples:

- `CD_IO_RQ_R_SM` is the number of requests to read small blocks on a cell disk.
- `GD_IO_BY_W_LG_SEC` is the number of MB of large block I/O per second on a grid disk.

Thresholds define extreme values for a metric, which might indicate a problem or other event of interest to an administrator. A user-defined threshold may be created or altered using the `CREATE THRESHOLD` or `ALTER THRESHOLD` CellCLI command. If a metric value crosses a user-defined threshold, an alert will be generated. A metric can be associated with warning threshold and a critical threshold.

For further details about Exadata cell metric and threshold attributes, refer to the *Oracle Exadata Storage Server Software User's Guide*.

Monitoring Exadata Cell Metrics: Examples

```
CellCLI> LIST METRICDEFINITION WHERE objectType ='CELL' DETAIL
         name: CL_CPUT
         description: "Cell CPU Utilization is the percentage of time over
                       the previous minute that the system CPUs were not
                       idle (from /proc/stat). "
         metricType: Instantaneous objectType: CELL
         unit: %
         ...
```

```
CellCLI> LIST METRICHISTORY WHERE name like 'CL_.*' -
         AND collectionTime > '2015-10-11T15:28:36-07:00'
         CL_RUNQ   cell03_2    6.0       2015-10-11T15:28:37-07:00
         CL_CPUT   cell03_2    47.6 %    2015-10-11T15:29:36-07:00
         CL_FANS   cell03_2    1         2015-10-11T15:29:36-07:00
         CL_TEMP   cell03_2    0.0 C     2015-10-11T15:29:36-07:00
         CL_RUNQ   cell03_2    5.2       2015-10-11T15:29:37-07:00
         ...
```

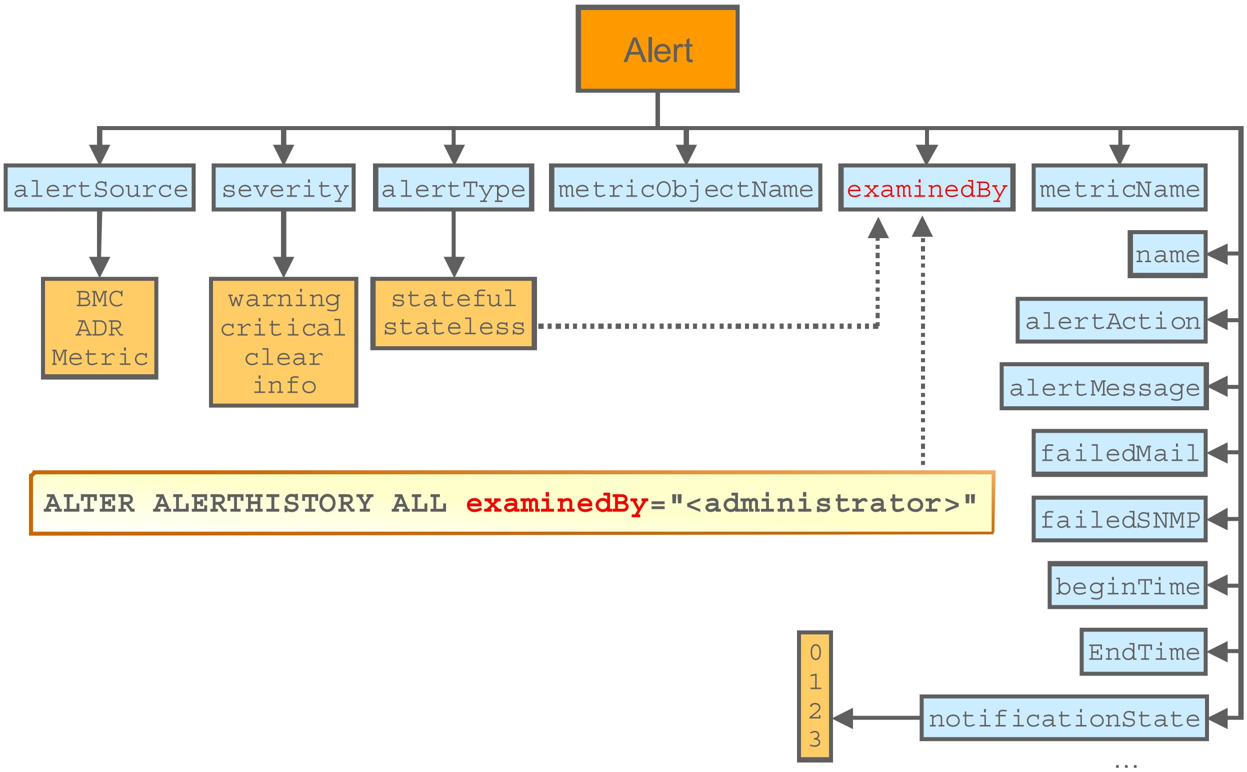
```
CellCLI> LIST METRICCURRENT WHERE objectType = 'CELLDISK'
         CD_IO_TM_W_SM_RQ CD_1_cell03 205.5 us/request
         CD_IO_TM_W_SM_RQ CD_2_cell03  93.3 us/request
         CD_IO_TM_W_SM_RQ CD_3_cell03   0.0 us/request
         ...
```

ORACLE

The slide shows some basic commands that you can use to display metric information:

- Use the `LIST METRICDEFINITION` command to display the metric definitions for the cell. A metric definition describes the configuration of a metric. The example in the slide does not specify any particular metric, so all metrics corresponding to the `WHERE` clause are printed. In addition to the `WHERE` clause, you can also specify the metric definition attributes that you want to print. If the `ATTRIBUTES` clause is not used, a default set of attributes is displayed. To list all the attributes, you can add the `DETAIL` keyword at the end of the command.
- Use the `LIST METRICHISTORY` command to display the metric history for the cell. A metric history describes a collection of past metric observations. Similar to the `LIST METRICDEFINITION` command, you can specify attribute filters, an attribute list, and the `DETAIL` keyword for the `LIST METRICHISTORY` command. The example in the slide lists metrics having names that start with `CL_` that were collected after the specified time.
- Use the `LIST METRICCURRENT` command to display the current metric values for the cell. The example in the slide lists all cell disk metrics. The metric values shown in the slide correspond to the average latency per request of writing small blocks to a cell disk. For this metric, there is a metric observation for every cell disk.

Monitoring Exadata Storage Server with Alerts



ORACLE

Exadata Storage Server alerts represent events of importance occurring within the storage cell, typically indicating that storage cell functionality is either compromised or in danger of failure. An administrator should investigate alerts, because they might require urgent corrective or preventative action. Notification of cell alerts by using email or SNMP can be configured during the initial configuration of Exadata. An existing configuration can be adjusted, or a new configuration can be established, at any time by using the `ALTER CELL` CellCLI command.

Cell alerts can be stateful or stateless. Stateful alerts represent observable cell states that can be subsequently retested to detect whether the state has changed, indicating that a previously observed alert condition is no longer a problem. Stateless alerts represent point-in-time events that do not represent a persistent condition; they simply show that something occurred.

Alerts can have one of the following severities: `warning`, `critical`, `clear`, or `info`. Examples of possible events that trigger alerts are physical disk failure, disk read/write errors, cell temperature exceeding recommended value, cell software failure, and excessive I/O latency.

Metrics can be used to signal stateful alerts by using `warning` or `critical` threshold values. When the metric value crosses the threshold value, an alert is signaled. An alert with a `clear` severity indicates that a previous critical or warning condition has returned to normal. For threshold-based alerts, a `clear` alert is generated when the measured value crosses back over the threshold value.

Alerts with an `info` severity are stateless and log conditions that might be informative to an administrator, but for which no administrator action is required. Informational alerts are not distributed by email or SNMP notifications.

The slide illustrates some of the important alert attributes:

- `name` provides an identifier for the alert.
- `alertSource` provides the source of the alert. Some possible sources are listed in the slide.
- `severity` determines the importance of the alert. Possible values are `warning`, `critical`, `clear`, and `info`.
- `alertType` provides the type of the alert: `stateful` or `stateless`. Stateful alerts are automatically cleared on transition to normal. Stateless alerts are never cleared unless you change the alert by setting the `examinedBy` attribute. This attribute identifies the administrator who reviewed the alert and is the only alert attribute that can be modified by the administrator by using the `ALTER ALERTHISTORY` command.
- `metricObjectName` is the object for which a metric threshold has caused an alert.
- `metricName` provides the metric name if the alert is based on a metric.
- `alertAction` is the recommended action to perform for this alert.
- `alertMessage` provides a brief explanation of the alert.
- `failedMail` is the intended email recipient when a notification failed.
- `failedSNMP` is the intended SNMP subscriber when a notification failed.
- `beginTime` provides the time stamp when an alert changes its state.
- `endTime` provides the time stamp for the end of the period when an alert changes its state.
- `notificationState` indicates progress in notifying subscribers to alert messages:
  - `0`: never tried
  - `1`: sent successfully
  - `2`: retrying (up to 5 times)
  - `3`: five failed retries

Note: Some I/O errors may result in an ASM disk going offline without generating a cell alert. You should continue to perform I/O monitoring from your databases and ASM environments to identify and remedy these kinds of problems.

Monitoring Cell Alerts and Creating Thresholds: Examples

```
CellCLI> LIST ALERTDEFINITION ATTRIBUTES name, metricName, description
    ADRAlert "CELL Incident Error"
    HardwareAlert "Hardware Alert"
    StatefulAlert_CG_IO_RQ_LG CG_IO_RQ_LG "Threshold Based Stateful Alert"
    StatefulAlert_CG_IO_RQ_LG_SEC CG_IO_RQ_LG_SEC "Threshold Based …Alert"
    StatefulAlert_CG_IO_RQ_SM CG_IO_RQ_SM "Threshold Based Stateful Alert"
    ...
```

```
CellCLI> LIST ALERTHISTORY WHERE severity = 'critical' -
         AND examinedBy = '' DETAIL
```

```
CellCLI> ALTER ALERTHISTORY 1671443814 examinedBy="JFV"
```

```
CellCLI> CREATE THRESHOLD ct_io_wt_lg_rq.interactive -
         warning=1000, critical=2000, comparison='>', -
         occurrences=2, observation=5
```

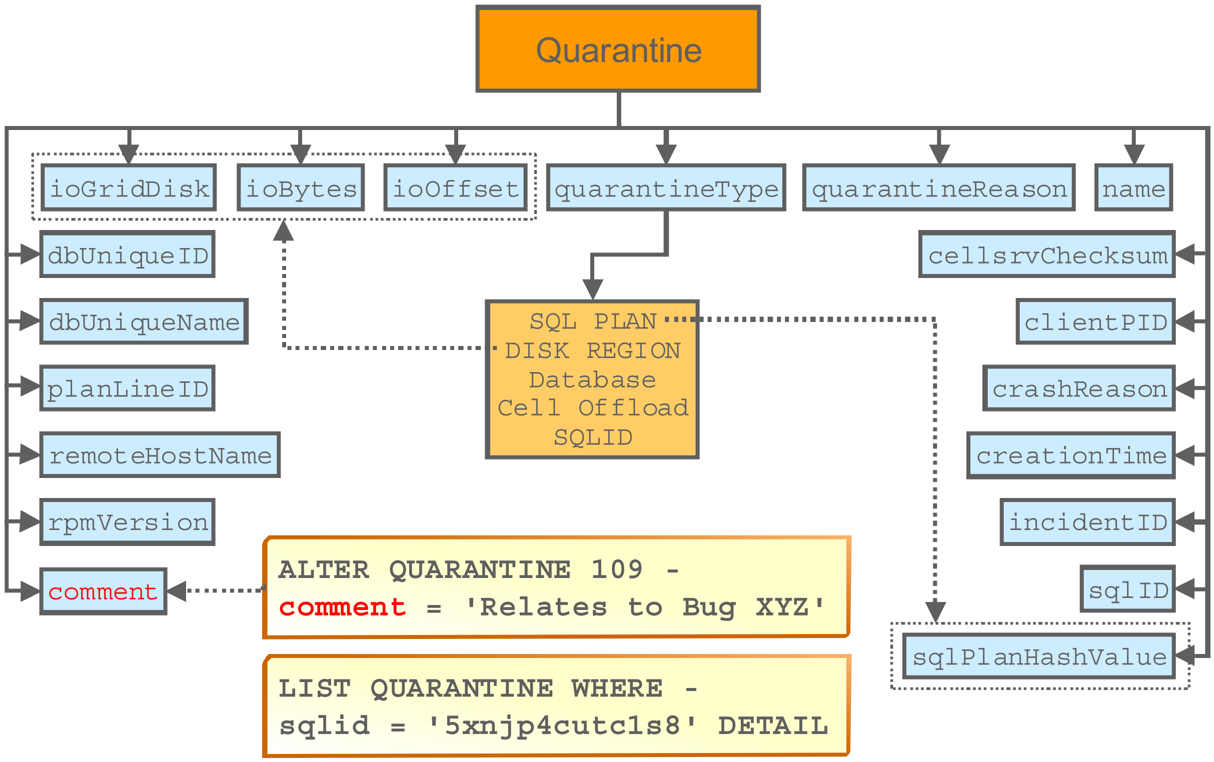
ORACLE

The slide shows some examples of CellCLI commands that display alert information. The following commands used for displaying alerts are very similar to the ones used for displaying metric information:

- **`LIST ALERTDEFINITION`:** Displays the definition for every alert that can be produced on the cell. The example in the slide displays the alert name, metric name, and description. The metric name identifies the metric on which the alert is based. `ADRAlert` and `HardwareAlert` are not based on any metric and, therefore, do not have metric names.
- **`LIST ALERTHISTORY`:** Displays the alert history for a cell. The example in the slide lists in detail all critical alerts that have not been reviewed by an administrator.
- **`ALTER ALERTHISTORY`:** Updates the alert history for the cell. The example in the slide shows how to set the `examinedBy` attribute to the user ID of the administrator that examined the alert. The `examinedBy` attribute is the only `ALERTHISTORY` attribute that can be modified. The example uses the alert sequence ID to identify the alert. `alertSequenceID` provides a unique sequence ID number for the alert. When an alert changes its state, another occurrence of the alert is created with the same sequence number, but with a different time stamp.

- **`CREATE THRESHOLD`:** Creates a threshold that specifies the conditions for the generation of a metric alert. The example creates a threshold for the `CT_IO_WT_LG_RQ` metric associated with the `INTERACTIVE` category. This metric specifies the average number of milliseconds that large I/O requests issued by the category have waited to be scheduled by IORM in the past minute. A large value indicates that the I/O workload from this category is exceeding the allocation specified for it in the IORM plan. The alert is triggered by two consecutive measurements (`occurrences=2`) over the threshold values: one second for a warning alert (`warning=1000`) and two seconds for a critical alert (`critical=2000`). The `observation` attribute is the number of measurements over which measured values are averaged.

Isolating Faults with Exadata Storage Server Quarantine



In addition to metrics and alerts, when prescribed faults are detected in Exadata Storage Server, a quarantine object is automatically created. By this, the action that caused the fault can be quarantined, so that the fault can be avoided in the future. Quarantine reduces the chance of storage server software crashes, and improves storage availability.

For example, if the cell crashes while performing Smart Scan for a SQL statement, Exadata Storage Server quarantines the SQL statement. Later, when the same SQL statement occurs again, the cell will not allow the SQL statement to use Smart Scan.

The following types of automatic quarantine are available:

- **`SQL PLAN`**: Created when the cell crashes while performing Smart Scan for a SQL statement. The SQL Plan for the SQL statement is quarantined, and Smart Scan is disabled for the SQL plan.
- **`DISK REGION`**: Created when the cell crashes while performing Smart Scan of a disk region. The 1 MB disk region being scanned is quarantined and Smart Scan is disabled for the disk region.
- **`Database`**: Created when the cell detects that a particular database causes instability. Instability detection is based on the number of SQL Plan Quarantines for a database. Smart Scan is disabled for the database.

- **`Cell Offload`:** Created when the cell detects that some offload feature has caused instability. Instability detection is based on the number of database quarantines for a cell. Smart Scan is disabled for all databases.

When a quarantine is created, an alert is generated to notify administrators about what was quarantined, why the quarantine was created, when and how the quarantine can be dropped manually, and when the quarantine is dropped automatically. All quarantines are automatically removed when a cell is patched or upgraded.

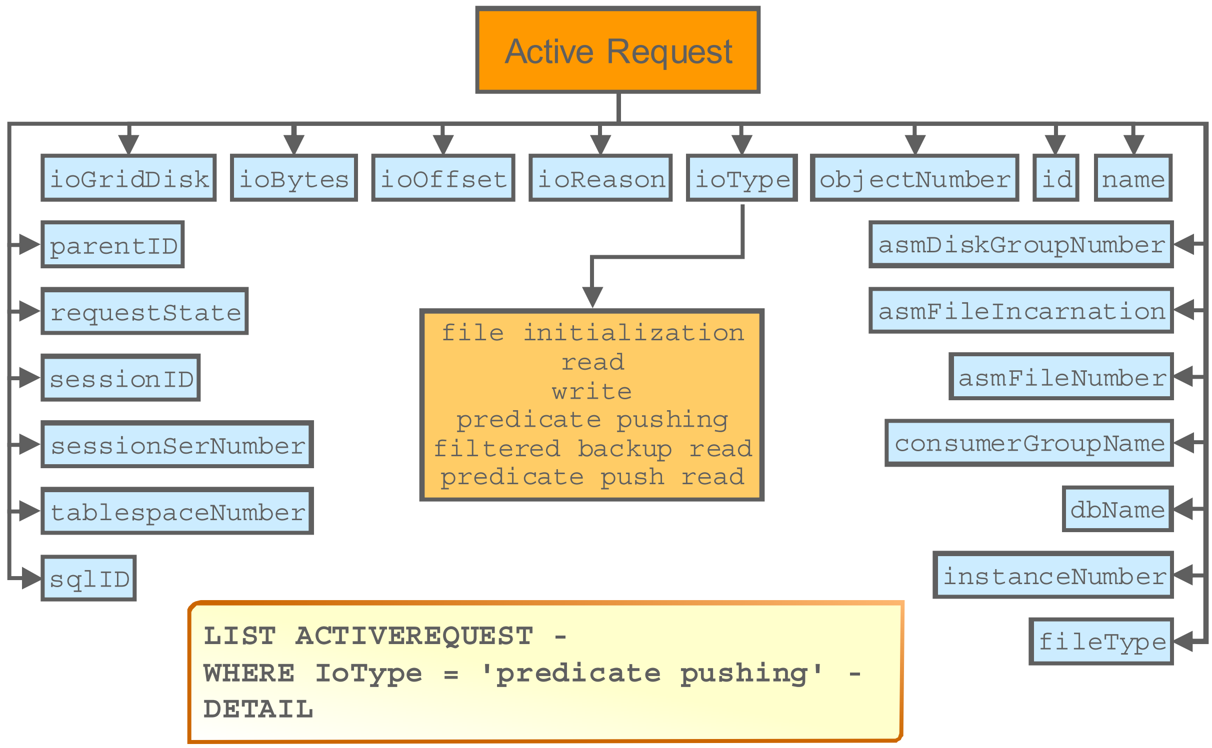
The Exadata Storage Server quarantine mechanism is designed to be automated and self-contained. However, the following CellCLI commands are available to manually manipulate quarantines:

- **`LIST QUARANTINE`:** To show the quarantines currently on the cell
- **`ALTER QUARANTINE`:** To set the comment attribute. The comment attribute is the only quarantine attribute that can be modified.
- **`DROP QUARANTINE`:** To manually remove a quarantine. In general, a quarantine can be removed if the quarantined entity is not expected to cause further problems. Refer to the alert message associated with the quarantine for more details before manually removing a quarantine.
- **`CREATE QUARANTINE`:** To manually create a quarantine object. Manual creation of quarantines should be done in coordination with Oracle Support Services. Manual quarantines are created to proactively isolate SQL statements that are known to cause problems. Following is an example of a manual quarantine creation:

```
CELLCLI> CREATE QUARANTINE  quarantineType="SQLID",
sqlid="5xnjp4cutc1s8"
```

The slide lists the attributes associated with the quarantine object. Note that the `sqlPLanHashValue` attribute is applicable only when the `quarantineType` is `SQL PLAN`. Likewise, the I/O attributes (`ioGridDisk`, `ioBytes` and `ioOffset`) are applicable only when the `quarantineType` is `DISK REGION`.

Monitoring Exadata Storage Server with Active Requests



An active request provides a client- or application-centric view of client I/O requests that are currently being processed by a cell.

The slide shows the most important attributes of an active request. You can see that an active request is characterized at all levels: instance, database, ASM, and cell. Most of the attributes have self-explanatory names. The following are some of the attributes:

- **`ioReason`**: Is the reason for the I/O activity, such as a control file read
- **`ioType`**: Identifies the type of active request. Possible values are listed in the slide.
- **`requestState`**: Identifies the state of the active request. Possible values include:
  - Accessing Disk
  - Network Receive
  - Queued Extent
  - Queued for File Initialization
  - Queued for Network Send
  - Queued for Read
  - Queued in Resource Manager
  - Computing Result
  - Network Send
  - Queued for Disk
  - Queued for Filtered Backup Read
  - Queued for Predicate Pushing
  - Queued for Write

Use the `LIST ACTIVEREQUEST` command to display active request details for the cell. The syntax is very similar to other `LIST` commands. You can specify which attributes to display or display them all by using the `DETAIL` clause. You can also filter the output by using a `WHERE` clause.

Automatic Hard Disk Scrubbing and Repair

- Disks are periodically scanned to detect latent corruptions
  - Scrubbing I/Os are issued when the disk is idle
- If a corruption is detected, and repair is required:
  - Data in the local Smart Flash Cache is used if possible
  - Otherwise, ASM is requested to read an available mirror
- Configuration options:
  - Setting the next start time:

```
CellCLI> ALTER CELL hardDiskScrubStartTime="<Timestamp>"
```

  - Setting the scrubbing interval:

```
CellCLI> ALTER CELL hardDiskScrubInterval = [ daily | weekly |
         biweekly | none ]
```

ORACLE

When data is read, checks are performed to validate its logical consistency. If a corruption is detected on `NORMAL` or `HIGH` redundancy disk groups, ASM can automatically recover by reading the mirrored data copies. One weakness with this approach is that corruption of seldom-accessed data could go unnoticed in the system for a long time between reads. Also, because reads are normally always directed to the primary data copy, corruption of secondary copies can go unnoticed while the primary copy is available. This can leave the system vulnerable to data loss if the primary copy also fails.

Exadata release 11.2.3.3.0, introduces a proactive hard disk scrubbing capability that minimizes the possibility of data loss due to latent corruptions on storage server hard disk drives.

Scrubbing works by periodically scanning each disk; every two weeks by default. So that scrubbing does not interfere with system performance, the scrubbing read I/Os are issued when the disk is idle. If a corruption is detected, the following occurs:

- If the data is dirty in write-back Smart Flash Cache, no further action is taken because the corruption will be overwritten when the cached copy is flushed.
- If the data is clean in Exadata Smart Flash Cache, the corruption is repaired by using the cached data.
- If a clean data copy is not in Exadata Smart Flash Cache, the corruption is repaired by using ASM to read an uncorrupted copy of the data.

Adaptive Hard Disk Scrubbing

- Disks with bad sectors are more likely to develop more bad sectors.
  - More frequent scrubbing makes sense.
- Starting with Exadata release 12.1.2.3.0, if a scrub detects bad sectors, the scrubbing interval may be automatically adjusted:
  - If the scrubbing interval is set to greater than one week, an additional scrub is scheduled in one week.

Proactive hard disk scrubbing automatically inspects and repairs hard disks periodically when the hard disks are idle. The default scrubbing schedule is every two weeks. However, when a hard disk starts to develop bad sectors, it is better to scrub that disk more frequently because it is likely to develop more bad sectors.

With Exadata release 12.1.2.3.0, if a bad sector is found on a hard disk in a current scrubbing job, Exadata will schedule a follow-up scrubbing job for that disk in one week. When no bad sectors are found in a scrubbing job for that disk, the schedule will fall back to the scrubbing schedule specified by the `hardDiskScrubInterval` attribute.

If the user has set the `hardDiskScrubInterval` to one week or less, Exadata will use the user-configured frequency instead of the weekly follow-up schedule even if bad sectors are found.

The minimum software that is required for this capability is Exadata Storage Server release 12.1.2.3.0 in conjunction with Oracle Grid Infrastructure version 11.2.0.4.16 (April 2015) or 12.1.0.2.4 (January 2015).

Cell Alert Summary

A periodic alert email summarizes the existing open alerts:

- Contains the most important information about open alerts
- Provides users with a concise summary of open issues without requiring cell access
- Makes it easier for users to identify open alerts
- Produced weekly, at 8am on Monday, by default
- Configuration options:
  - Setting the next delivery time:

```
CellCLI> ALTER CELL alertSummaryStartTime="<Timestamp>"
```

  - Setting the delivery interval:

```
CellCLI> ALTER CELL alertSummaryInterval = [ daily | weekly |
         biweekly | none ]
```


Exadata release 11.2.3.3.0 introduces a new cell alert summary, which periodically summarize the most important information about open alerts into a HTML document.

The periodic alert email includes a table of open alerts containing the alert name, event time, severity, and alert description, making it quick and easy for users to identify the most important issues.

Previously, to manually identify open alerts, users needed to connect to the cell and examine the alert history, which contains both resolved and unresolved alerts. The cell alert summary provides users a concise summary of all the open issues on the cell, without requiring access to the cell.

Cell Diagnostic Packages

A cell diagnostic package contains related logs and traces:

- A diagnostic package is generated for each cell alert.
  - `LIST DIAGPACK` lists previously generated packages.
- Use `CREATE DIAGPACK` to manually create a package:

```
CellCLI> CREATE DIAGPACK packStartTime="<Timestamp>"
           [, durationInHrs=<duration> ]
```

```
CellCLI> CREATE DIAGPACK alertName=<alertName>
```

- Use the following methods to access diagnostic packages:
  - File location: Look under `$LOG_HOME` on each cell.
  - Email alerts: See attachment for each email alert.
  - Web interface: `https://<cellHostname>/diagpack`
  - REST APIs: `https://<cellHostname>/diagpack/download?name=<packName>`
    `https://<cellHostname>/diagpack/download?alert=<alertName>`
  - ExaCLI: `exacli> DOWNLOAD DIAGPACK <packName> <destinationFolder>`

ORACLE

Starting with Exadata release 12.1.2.2.0, cells automatically generate customized diagnostic packages that include relevant logs and traces upon generating an alert. This applies to all cell alerts, including both hardware alerts and software alerts. The timely collection of diagnostic information prevents rollover of critical logs.

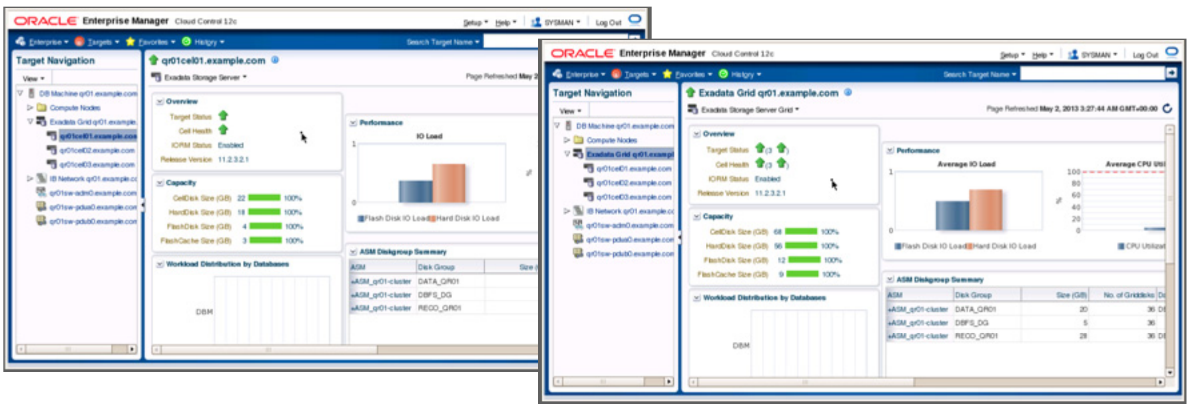
Cell administrators can display a list of previously generated diagnostic packages by using the `LIST DIAGPACK` CellCLI command. The `CREATE DIAGPACK` command enables administrators to create custom diagnostic packages relating to a specified time period. Starting with Exadata release 12.1.2.3.0, the `CREATE DIAGPACK` command also supports re-creating diagnostic packages for a specified alert by using the `alertName` parameter.

Diagnostic packages are stored in `$LOG_HOME` on each storage server and MS sends the diagnostic package as an email attachment for every email alert. In addition, appropriately privileged cell administrators can access cell diagnostic packages by using a web interface, by calling a REST API, or by using the ExaCLI `DOWNLOAD DIAGPACK` command.

See `CREATE DIAGPACK` in the *Oracle Exadata Storage Server Software User's Guide* for details.

Monitoring Exadata Storage Server with Enterprise Manager: Overview

- Each storage server is a separate target in Enterprise Manager.
  - Storage servers in a Database Machine are grouped together.
- Metrics in Enterprise Manager are mostly based on cell metrics.
- Additional thresholds can be set in Enterprise Manager.
- Alerts generated in the cell are displayed in Enterprise Manager.
- Additional alerts can be generated in Enterprise Manager.



ORACLE

Each Exadata Storage Server is monitored as a separate Enterprise Manager target, which covers the hardware, operating system, and Exadata Storage Server software. In addition, storage servers in a Database Machine are grouped together as an Exadata Storage Server Grid (or Exadata Grid) target. An Exadata Grid target facilitates collective monitoring of Exadata storage in Enterprise Manager Cloud Control.

Enterprise Manager in conjunction with the Exadata Storage Server plug-in leverages and extends the system of metrics and alerts within each Exadata Storage Server.

When Enterprise Manager is used to monitor storage servers, there are two different types of related metrics: storage server metrics and Enterprise Manager metrics. In most cases, there is a one-to-one mapping between them. The Management Server (MS) on Exadata Storage Server collects, computes, and manages metrics inside each individual storage server, and Enterprise Manager gathers and presents metrics from all the cells it monitors.

In addition to cell-level metric thresholds, you can set separate Enterprise Manager thresholds. In Exadata Storage Server software releases before release 11.2.2.3.1, cell thresholds can be set only on a limited set of cell metrics. Therefore, Enterprise Manager thresholds offer the opportunity to automatically raise alerts based on a wider range of metrics.

When equivalent metrics can be defined in either Enterprise Manager or at the cell level, the general recommendation is to define the threshold at the cell level so that the threshold definition and any associated alerts are maintained within the cell and are also visible using Enterprise Manager.

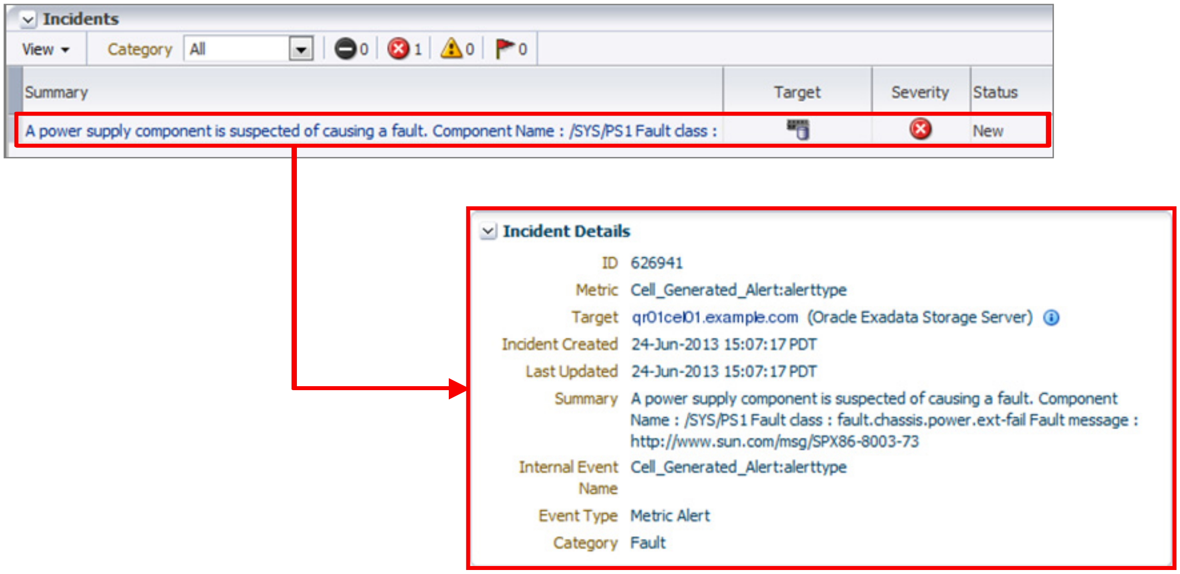
All Exadata Storage Server alerts are delivered by the storage server to Enterprise Manager by using Simple Network Management Protocol (SNMP). There are two types of server alerts that come from Exadata Storage Server:

- For Integrated Lights Out Manager (ILOM)–monitored hardware components, ILOM reports a failure or threshold exceeded condition as an SNMP trap, which is received by MS. MS processes the trap, creates an alert for the storage server, and delivers the alert via SNMP to Enterprise Manager.
- For MS–monitored hardware and software components, MS processes the failure or threshold exceeded condition, creates an alert, and delivers the alert via SNMP to Enterprise Manager.

From an end-user perspective, there is no difference between these two kinds of alerts.

An alert message may contain a corrective action, which must be performed to resolve the alert.

Monitoring Hardware Failure and Sensor State




```
CellCLI> list alerthistory
```

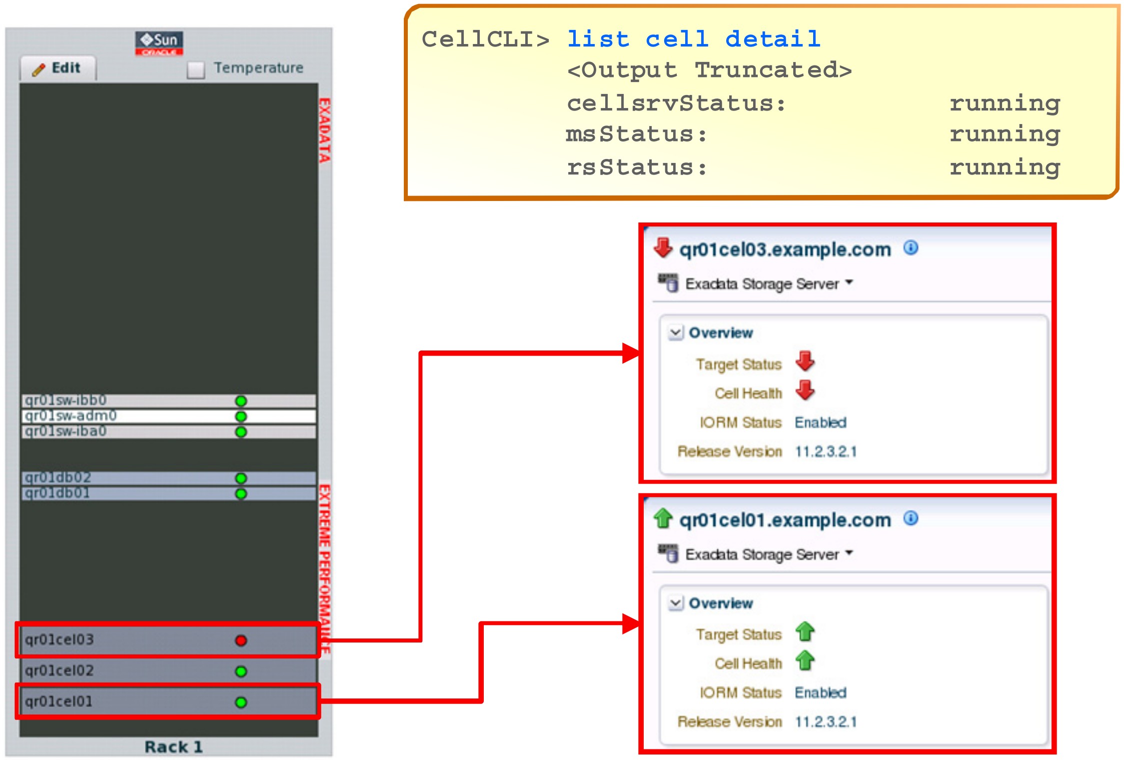
ORACLE

The hardware of an Exadata Storage Server is monitored collectively by ILOM and MS. ILOM monitors availability and sensor state by using preset thresholds for Exadata Storage Server hardware components, such as the system motherboard, processors, memory, power supplies, fans, and network interface controllers. MS monitors other hardware components directly, including the disk controller, hard disk drives, flash accelerator cards, and InfiniBand host channel adapter (HCA).

Together ILOM and MS provide full hardware monitoring and alerting. When an issue arises, an alert is automatically generated within the cell, and is also propagated to Enterprise Manager (if configured). The screenshots in the slide shows an example of an alert notification for a hardware fault.

If Enterprise Manager or some other SNMP manager is not configured, administrators must configure cells to deliver alerts by using email or they should periodically check the cell by using the CellCLI `LIST ALERTHISTORY` command.

Monitoring Exadata Storage Server Availability



ORACLE

Monitoring of Exadata Storage Server availability is built into Enterprise Manager. The screenshots in the slide show examples of how Enterprise Manager indicates that storage servers are available or unavailable.

For environments without Enterprise Manager, storage server availability should be checked by confirming network connectivity. In addition, the status of the cell services should be checked by using the CellCLI `LIST CELL DETAIL` command. The output for a healthy cell should contain the following:

```
cellsrvStatus:          running
msStatus:               running
rsStatus:               running
```

Checking for Undelivered Alerts

- Periodically check for undelivered alerts.
  - CellCLI example that checks one cell:

```
CellCLI> list alerthistory where notificationState != 1 and examinedBy = ''
         1_1      2015-04-25T12:09:22-07:00       warning         "The warning
    threshold for the following metric has been crossed. Metric Name : CL_MEMUT
    Metric Description : Percentage of total physical memory on the cell that is
    currently used  Object Name : exa1cel01  Current Value : 51.0 %  Threshold Value
    : 50.0 %  "

CellCLI>
```

  - `dcli` example that checks all the cells listed in the `cell_group` file:

```
$ dcli -g cell_group cellcli -e "list alerthistory where notificationState != 1
  and examinedBy = \'\'"
```

- If undelivered alerts exist, check cell-to-agent network connectivity, agent availability, and cell configuration.

ORACLE

Relying on Enterprise Manager for alert notifications depends on alerts being reliably propagated to Enterprise Manager. Although this is usually the case, administrators should periodically check for undelivered alerts.

The CellCLI `LIST ALERTHISTORY` command can be used to check for undelivered alerts. The first example in the slide shows how to check for alerts that do not have a success status (`1`) and that have not been marked as examined by an administrator. The second example shows how to use the same command in conjunction with `dcli` to check a group of cells at once.

Note that an alert might not have a success state for many reasons and that the output generated by the monitoring command in the slide does not necessarily relate to a cell fault or misconfiguration. For example, an alert might not be propagated because the agent is unavailable. Also, if the cell is configured to deliver alerts by using SNMP and email, it is possible for the alert to show an unsuccessful notification state when one delivery method succeeds but the other one fails. Finally, if communication is disrupted between a storage server and the agent where the Exadata Storage Server plug-in is deployed, alerts processed by MS may not be delivered to Enterprise Manager. Remember that MS will retry up to five times to deliver the alert, so if the communication disruption is temporary, you may observe an undelivered alert and then see the same alert delivered moments later.

Checking for Disk I/O Errors

To check for disk I/O errors, create the following warning threshold on every cell:

```
CellCLI> create threshold CD_IO_ERRS_MIN comparison='>', warning=0, -
> occurrences=1, observation=1
Threshold CD_IO_ERRS_MIN successfully created

CellCLI> list threshold CD_IO_ERRS_MIN detail
         name:                   CD_IO_ERRS_MIN
         comparison:             >
         observation:            1
         occurrences:            1
         warning:                0.0

CellCLI>
```

ORACLE

The slide shows a recommended storage server threshold setting that assists in checking for disk I/O errors.

Disk I/O errors are typically handled by Exadata cell software automatically. The recommended threshold setting causes an alert to be generated by the cell when an I/O error is detected.

No action is required for a single warning alert. However many warnings for a single drive may by a precursor to a drive failure. Likewise, many warnings across all the drives in a single cell may indicate an impending fault with a controller or some other component.

If this warning is reported and a critical alert is also generated by MS for a disk component, follow the action specified in the critical alert.

Checking for Network Errors

To monitor the storage network for dropped packets, create the following warning thresholds on every cell:

```
CellCLI> create threshold N_MB_DROP_SEC comparison='>', warning=0, -
> occurrences=1, observation=1
```

```
CellCLI> create threshold N_MB_RDMA_DROP_SEC comparison='>', warning=0, -
> occurrences=1, observation=1
```

ORACLE

A functioning storage network is essential to the overall health and performance of Exadata. A loss of functionality is often indicated by the amount of dropped network traffic. Normally, dropped network traffic is automatically handled within Exadata, so occasional dropped packets are not a cause for concern. However, many dropped packets for a single cell may be a precursor to a problem.

Periodic monitoring of the `N_MB_DROP_SEC` and `N_MB_RDMA_DROP_SEC` cell metrics is advised and the slide shows two recommended threshold settings that can be used to assist. Note that the thresholds must be defined in each storage server.

Monitoring File System Free Space

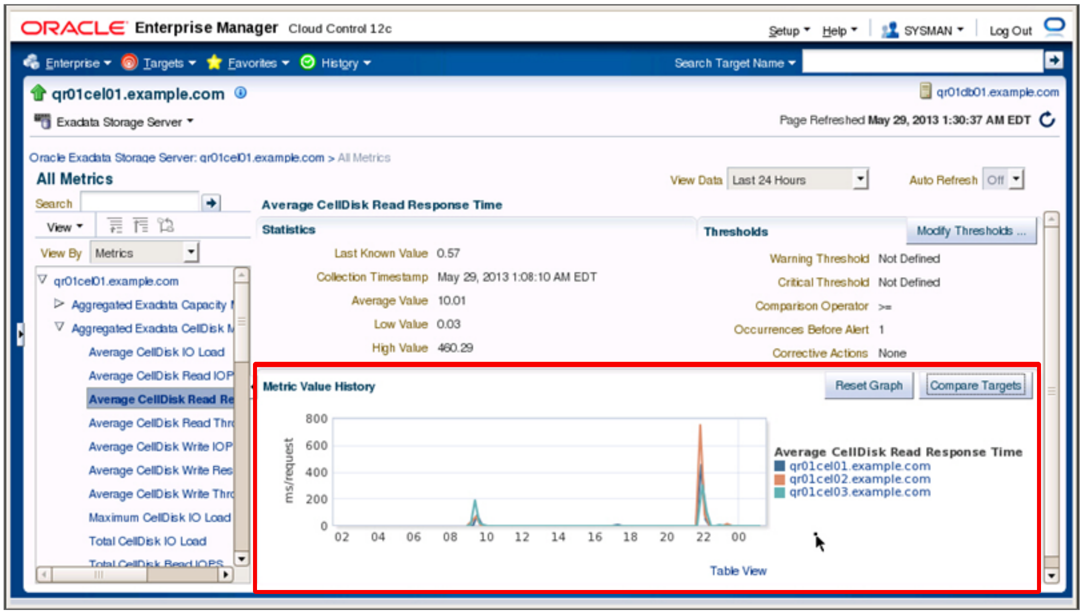
File system free space on Exadata Storage Servers is monitored automatically by MS:

- No direct administrator action is required.
- MS generates an alert if free space becomes low.
- MS automatically reclaims used space by purging old log files, trace files, crash dumps, and other unnecessary files.

The Exadata Storage Server software automatically monitors file system utilization on each storage server. There is no need to set a metric threshold for file system utilization within the cell or in Enterprise Manager. MS monitors file system free space, generates critical alerts when free space becomes low, and takes corrective action to free used space. For detailed information, see the section titled "Understanding Automated Cell Maintenance" in the *Oracle Exadata Storage Server Software User's Guide 11g Release 2 (11.2)*.

Comparing Metrics Across Multiple Storage Servers

- Enterprise Manager can compare metrics across storage servers.
- Large imbalances may indicate a problem.



ORACLE

Exactly how data is spread across a Database Machine depends upon many factors, including the number of applications and databases that are consolidated on the environment, along with how storage is configured to meet various performance, availability, security, and other requirements.

However, in general terms, work performed by storage servers to satisfy database requests should be evenly spread across multiple storage servers. A significant imbalance in resource utilization or I/O performance for a single storage server compared to the others should be a cause for further investigation.

Various CPU, memory, and I/O utilization metrics can be periodically compared across the storage servers to ensure that there are no unwarranted imbalances. Enterprise Manager users can use the Compare Targets button on the All Metrics page to quickly and easily perform metric comparisons across a group of storage servers. The screenshot in the slide shows an example comparing the Average CellDisk Read Response Time metric on three cells. The chart shows similar metrics peaks over time indicating a well-balanced system.

Monitoring Metrics in a Storage Server

- There are no default metric thresholds for CPU, memory, or I/O utilization.
- 100 percent utilization may be normal in certain cases.
- Metric observations are useful and relevant when compared with previously captured baselines.
- Suggested methodology:
  1. Determine key metrics based on your applications and service requirements.
  2. Record baseline metric observations based on normal and peak system usage.
  3. Implement thresholds to generate alerts for observations outside expected ranges.
     - Warning thresholds to indicate slightly abnormal observations
     - Critical thresholds to indicate situations that could affect service levels

Exadata Storage Server is designed to be a well-balanced storage subsystem. As a result, CPU, memory, and I/O metric values on Exadata Storage Server can be higher than a typical database or general-purpose server, but still be in proper operating range.

For example, CPU may be 100% utilized while performing queries against data using Hybrid Columnar Compression (HCC). In this case, 100% CPU utilization is not problematic and no alert should be generated.

Resource usage metric observations within a storage server are useful and relevant only if they are compared with values captured during a period of normal operation. The slide outlines a suggested methodology to determine important metrics and implement thresholds based on previously recorded observations.

Third-Party Monitoring Tools

- Installing additional software, including any third-party monitoring agent, is not supported on Exadata Storage Server.
- Exadata Storage Server can be configured to send alerts to any SNMP subscriber.

It is not permissible to install any additional software, including third-party monitoring agents, on Exadata Storage Server.

It is permissible to configure Exadata Storage Server to send alerts by using SNMP to third-party management consoles.

Quiz

Where should you define thresholds when using Enterprise Manager to monitor Exadata Storage Servers?

a. In Enterprise Manager only
b. Inside the Exadata cell only
c. Inside the Exadata cell, if possible. Otherwise, in Enterprise Manager.
d. Either in Enterprise Manager or inside the Exadata cell. It does not matter.

Answer: c

When equivalent metrics can be defined in either Enterprise Manager or at the cell level, the general recommendation is to define the threshold at the cell level so that the threshold definition and any associated alerts are maintained within the cell.

Quiz

Assuming that Enterprise Manager is properly configured to monitor Exadata Storage Server, which of the following can occur?

a. A metric observation generates an alert in the cell, but no alert is seen in Enterprise Manager.
b. A metric observation generates an alert in the cell and an alert is seen in Enterprise Manager.
c. A metric observation generates no cell alert, but an alert is seen in Enterprise Manager.
d. A metric observation results in two alerts in Enterprise Manager.

Answer: a, b, c, d

A metric observation generates an alert in the cell, but no alert is seen in Enterprise Manager: This can occur when a cell-based threshold triggers a cell alert, but the alert is not propagated to Enterprise Manager due to a network communication problem.

A metric observation generates an alert in the cell and an alert is seen in Enterprise Manager: This is normal and an expected behavior. It occurs when a cell-based threshold triggers an alert and the alert is propagated to Enterprise Manager.

A metric observation generates no cell alert, but an alert is seen in Enterprise Manager: This occurs when a threshold set in Enterprise Manager triggers an alert, and there is either no cell-level threshold for the metric or the cell-level threshold is set at a level that does not trigger an alert in the cell.

A metric observation results in two alerts in Enterprise Manager: This occurs when a cell-level threshold and an Enterprise Manager threshold are defined on the same metric, and the corresponding cell-level alert is propagated to Enterprise Manager.

Summary

In this lesson, you should have learned how to:

- Describe Exadata Storage Server metrics, alerts, and active requests
- Identify the recommended focus areas for Exadata Storage Server monitoring
- Describe how to monitor the recommended Exadata Storage Server focus areas

Practice 15 Overview: Monitoring Exadata Storage Server

In these practices, you will:

- Monitor Exadata Storage Server using metrics, alerts, and active requests
- Perform Exadata Storage Server monitoring by using Enterprise Manager

16

Monitoring Exadata Database Machine Database Servers

Objectives

After completing this lesson, you should be able to describe the monitoring recommendations for Exadata database servers.

Monitoring Database Servers: Overview

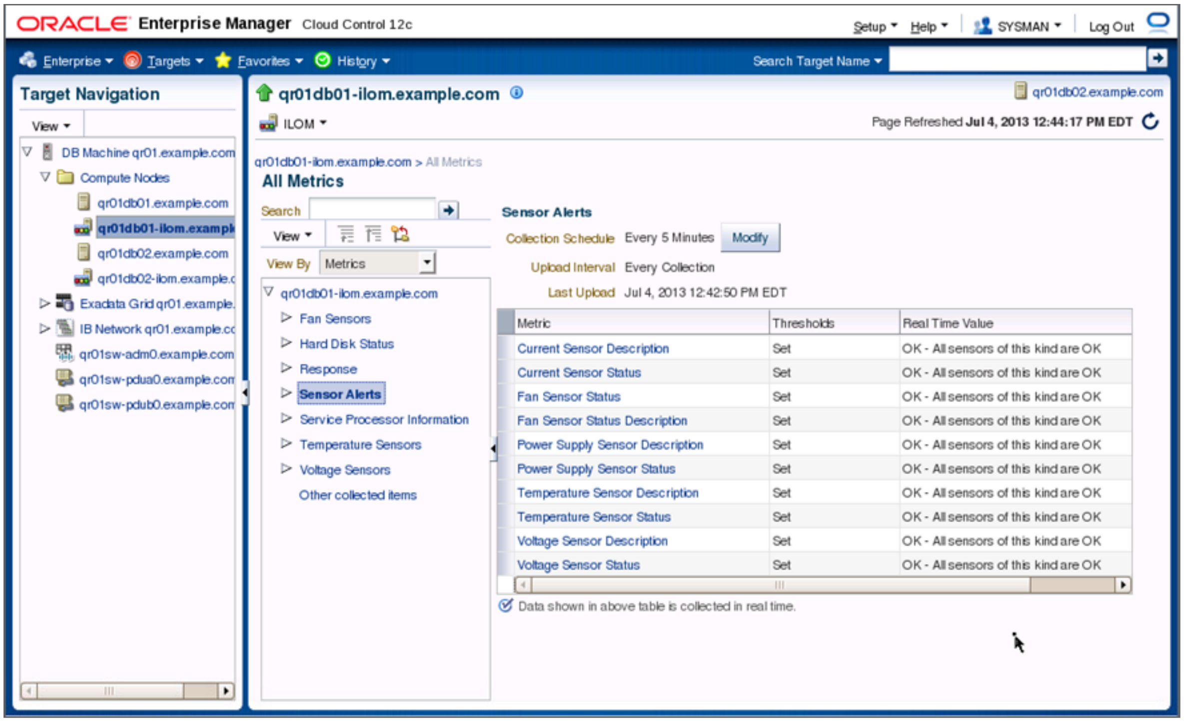
- Exadata database server monitoring is essentially the same as monitoring any clustered Oracle Database server.
  - Skills and practices are readily transferrable.
- This lesson covers Exadata–specific differences and recommendations with a focus on using Enterprise Manager to monitor:
  - Database server hardware
  - Database server operating system
  - Oracle Grid Infrastructure
  - Oracle Database
  - Oracle Management Agent

ORACLE

For the most part, monitoring Exadata database servers is the same as monitoring any other clustered Oracle Database server. So the skills and practices already used by Oracle database administrators are readily transferrable to an Exadata environment.

The specific differences and recommendations associated with Exadata and Enterprise Manager are the focus of this lesson.

Monitoring Hardware



Integrated Lights Out Manager (ILOM) monitors availability and sensor state using preset thresholds for database server hardware components, such as the system motherboard, processors, memory, power supplies, fans, and network interface controllers. The availability and sensor state can be monitored in Enterprise Manager. The following recommendations apply:

- There are no Exadata-specific thresholds to set for database server hardware monitoring. Failure conditions and threshold settings for the components monitored by ILOM are preset and are sufficient for the necessary level of monitoring.
- To view current sensor readings, log in to the Enterprise Manager Console and navigate to All Metrics from the ILOM target home page. To view current component status, including those that have a Faulted status, expand the Sensor Alerts entry in the hierarchical list and review the metrics for each sensor. An example is shown in the screenshot in the slide.
- To view the history of sensor state notifications generated by a sensor, click the link associated with the sensor.
- Components with a Faulted status will generate an alert. Any active alert will be visible on the home page of the ILOM target.

Monitoring the Operating System

- Operating system monitoring is built into Enterprise Manager by using the Host target type:
  - There are no Exadata-specific configuration requirements for general monitoring.
  - Thresholds can be set or changed to accommodate site-specific requirements.
  - Essential operating system alerts are generated by Enterprise Manager based on default metric thresholds.
- Monitor disk I/O in the storage servers and not on the database servers.

ORACLE

The database server operating system is viewed in Enterprise Manager as a Host target. The following recommendations apply:

- There are no Exadata-specific configuration requirements. The metrics and default thresholds provided by Enterprise Manager are sufficient for general monitoring.
- Thresholds can be set or changed on the Metric and Collection Settings page in Enterprise Manager to handle site-specific requirements.
- Database server operating system alerts are generated automatically by Enterprise Manager based on the default metric thresholds for the Host target, along with any user-defined threshold settings.
- Remember that I/O against Oracle Database data files is directed to Exadata Storage Servers by using the InfiniBand network. Therefore, metrics on database servers relating to disk I/O cannot be used to monitor Oracle Database data files. I/O relating to Oracle Database data files should be monitored on Exadata Storage Servers.

Monitoring Oracle Grid Infrastructure and Database

- Grid Infrastructure and Oracle Database monitoring is built into Enterprise Manager.
  - There are no Exadata-specific configuration requirements for general monitoring.
  - Thresholds can be set or changed to accommodate site-specific requirements.
  - Essential alerts are generated by Enterprise Manager based on default metric thresholds.

ORACLE

Oracle Automatic Storage Management (ASM), listener, and clusterware are viewed in Enterprise Manager as ASM, Listener, and Cluster targets, respectively. Oracle databases are viewed in Enterprise Manager as Cluster Database and Database Instance targets. The following recommendations apply:

- There are no Exadata-specific configuration requirements. The metrics and default thresholds provided by Enterprise Manager are sufficient for general monitoring.
- Thresholds can be set or changed on the Metric and Collection Settings page in Enterprise Manager to handle site-specific requirements. For example, thresholds are typically configured to provide a warning when the amount of free space in a disk group (`USABLE_FREE_MB`) falls below the required amount of free space to maintain redundancy if a failure group is lost (`REQUIRED_MIRROR_FREE_MB`).
- Alerts are generated automatically by Enterprise Manager based on default metric thresholds for the ASM, Listener, and Cluster targets, along with any user-defined threshold settings.

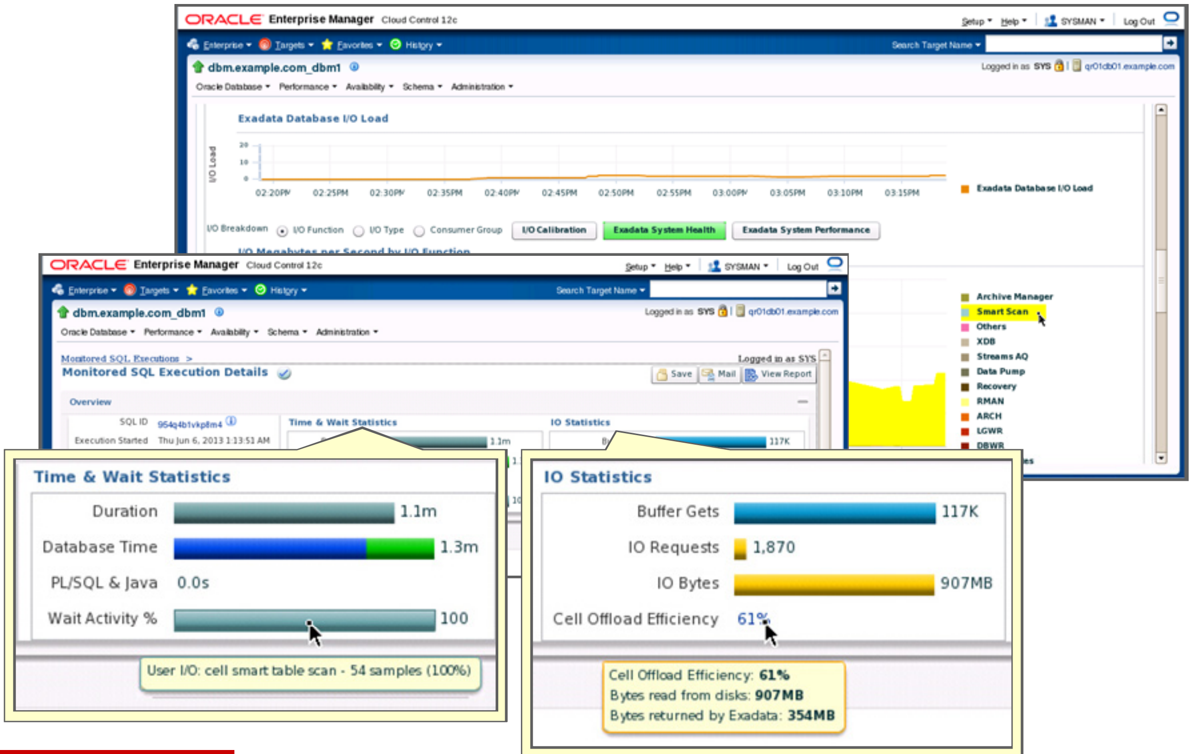
Monitoring Oracle Management Agent

- Oracle Management Agent monitoring is built in to Enterprise Manager by using the Agent target type.
  - There are no Exadata-specific configuration requirements for general monitoring.
  - Thresholds can be set or changed to accommodate site-specific requirements.
  - Essential alerts are generated by Enterprise Manager based on default metric thresholds.
- Remember that agent availability is especially important for the agents monitoring the Exadata targets.

Oracle Management Agents are viewed in Enterprise Manager as agent targets. The following recommendations apply:

- There are no Exadata-specific configuration requirements. The metrics and default thresholds provided by Enterprise Manager are sufficient for general monitoring.
- Thresholds can be set or changed on the Metric and Collection Settings page in Enterprise Manager to handle site-specific requirements.
- Alerts are generated automatically by Enterprise Manager based on default metric thresholds. By default, Enterprise Manager generates an alert if an agent is down or loses contact with the Management Server.
- Agents on your database servers monitor other Exadata components, such as the Exadata Storage Servers and the InfiniBand network, and the ability to monitor many Exadata components relies on the continued availability of the agent.

Database Monitoring with Enterprise Manager Cloud Control



Database monitoring by using Enterprise Manager Cloud Control for databases hosted on Exadata is essentially the same as for databases hosted on all other platforms. However, Enterprise Manager Cloud Control does provide Exadata-specific monitoring information. This includes access to Exadata-specific statistics and wait events, along with Exadata-specific enhancements to the database performance monitoring interfaces. The screenshots in the slide show some examples.

Monitoring Database Servers with MS and DBMCLI: Overview

- Exadata database servers now have a management server (MS) similar to Exadata Storage Servers.
- Monitoring and administration is performed using DBMCLI:
  - DBMCLI works in conjunction with MS to perform monitoring and administration tasks.
  - DBMCLI session example:

```
[dbmadmin@exa1db01 ~]$ dbmcli
DBMCLI: Release 12.1.2.1.0 - Production ...

DBMCLI> list dbserver
     exa1db01   online

DBMCLI> exit
quitting

[dbmadmin@exa1db01 ~]$
```


Commencing with Exadata software release 12.1.2.1.0, Exadata database servers contain a management server (MS) and restart server (RS), similar to the MS and RS on Exadata Storage Servers. On Exadata database servers, MS provides the following:

- Hardware and software monitoring including monitoring of hard disks, CPUs, and InfiniBand ports.
- System metric collection and monitoring.
- Alerting capabilities with the ability to create user-defined thresholds.

A command-line interface called DBMCLI works in conjunction with MS to configure, monitor, and manage the database servers. DBMCLI also configures Auto Service Request (ASR), capacity-on-demand, and database server email alerts.

The DBMCLI command set is essentially a small subset of the commands available in CellCLI, and most DBMCLI commands are similar to their equivalent CellCLI command.

Running DBMCLI

- Each Exadata database server is configured with operating system user accounts for use in conjunction with DBMCLI:
  - The `dbmadmin` user can:
    - Perform administrative tasks (`CREATE`, `DROP`, `ALTER`) using the DBMCLI utility
    - The `root` user can also do everything that `dbmadmin` can
  - The `dbmmonitor` user can only view (`LIST`, `DESCRIBE`) database server objects using the DBMCLI utility.
- To run DBMCLI, you can:
  - Execute DBMCLI from the OS command line
  - Execute DBMCLI remotely using SSH
  - Execute multiple DBMCLI commands simultaneously using `dcli`

ORACLE

Two operating system users are configured on each Exadata database server for use in conjunction with DBMCLI: `dbmadmin`, and `dbmmonitor`. The slide describes the function of each user account. The `root` operating system administrator can also execute all of the commands that are available to the `dbmadmin` user.

Like CellCLI, you can start DBMCLI from the operating system command line on the database server that you want to manage, or you can run it remotely from a network-attached client using Secure Shell (SSH). You can also use the `dcli` utility to execute a DBMCLI command simultaneously on multiple Exadata database servers.

Starting and Stopping Management Services on Exadata Database Servers

```
[dbmadmin@exa1db01 ~]$ dbmcli
DBMCLI: Release 12.1.2.1.0 ...

DBMCLI> ALTER DBSERVER RESTART SERVICES ALL

Stopping the RS and MS services...
The SHUTDOWN of services was successful.
Starting the RS and MS services...
Getting the state of RS services...  running
Starting MS services...
The STARTUP of MS services was successful.

DBMCLI>
```

ORACLE

Occasionally, you may need to start, stop, or restart the Exadata database server services: Restart Server (RS) and Management Server (MS). This can be achieved by using the DBMCLI `ALTER DBSERVER` command. An example is shown on the slide. Following is a summary of the relevant syntax:

```
ALTER DBSERVER { SHUTDOWN | RESTART | STARTUP } SERVICES { RS | MS | ALL}
```

Configuring Management Services on Exadata Database Servers

```
[dbmadmin@exa1db01 ~]$ dbmcli
DBMCLI: Release 12.1.2.1.0 ...

DBMCLI> ALTER DBSERVER smtpServer='my_mail.example.com',          -
                       smtpFromAddr='exadata.exa1db01@example.com',  -
                       smtpPwd=<email_address_password>             -
                       smtpToAddr='jane.smith@example.com',         -
                       notificationPolicy='critical,warning,clear',  -
                       notificationMethod='mail'
DBServer exa1db01 successfully altered

DBMCLI>
```

ORACLE

You can change or configure cell attributes by using the DBMCLI `ALTER DBSERVER` command.

The slide shows an example `ALTER DBSERVER` command that configures email notification. This facility sends email messages to the administrator (`jane.smith@example.com`) of the database server whenever critical, warning, and clear alerts are detected in the server. In addition to email notification, it is possible to configure notification by using Simple Network Management Protocol (SNMP).

Monitoring Database Server Metrics: Examples

```
DBMCLI> LIST METRICDEFINITION WHERE objectType ='DBSERVER' DETAIL
        ...
        name:                   N_NIC_KB_TRANS_SEC
        description:            "Number of kilobytes transmitted by
Ethernet interfaces per second"
        metricType:             Rate
        objectType:             DBSERVER
        unit:                   KB/sec
```

```
DBMCLI> LIST METRICHISTORY WHERE name like '.*TRANS.*' -
        AND collectionTime > '2015-12-01T15:28:36-08:00'
     N_HCA_MB_TRANS_SEC   exa1db01  0.242 MB/sec  2015-12-01T15:29:28-08:00
     N_IB_MB_TRANS_SEC    HCA-1:1   0.095 MB/sec  2015-12-01T15:29:28-08:00
     N_IB_MB_TRANS_SEC    HCA-1:2   0.147 MB/sec  2015-12-01T15:29:28-08:00
     N_IB_UTIL_TRANS      HCA-1:1   0.0 %         2015-12-01T15:29:28-08:00
     N_IB_UTIL_TRANS      HCA-1:2   0.0 %         2015-12-01T15:29:28-08:00
     ...
```

```
DBMCLI> LIST METRICCURRENT WHERE objectType = 'IBPORT'
        N_IB_MB_RCV_SEC         HCA-1:1         0.148 MB/sec
        N_IB_MB_RCV_SEC         HCA-1:2         0.894 MB/sec
        N_IB_MB_TRANS_SEC       HCA-1:1         0.373 MB/sec
        ...
```

ORACLE

The slide shows some basic commands that you can use to display metric information:

- Use the `LIST METRICDEFINITION` command to display metric definitions. A metric definition describes the configuration of a metric. The example in the slide does not specify any particular metric, so all metrics corresponding to the `WHERE` clause are printed. In addition to the `WHERE` clause, you can also specify the metric definition attributes that you want to print. If the `ATTRIBUTES` clause is not used, a default set of attributes is displayed. To list all the attributes, you can add the `DETAIL` keyword at the end of the command.
- Use the `LIST METRICHISTORY` command to display metric history. A metric history describes a collection of past metric observations. Similar to the `LIST METRICDEFINITION` command, you can specify attribute filters, an attribute list, and the `DETAIL` keyword for the `LIST METRICHISTORY` command. The example in the slide lists metrics having names that contain `TRANS` that were collected after the specified time.
- Use the `LIST METRICCURRENT` command to display current metric values. The example in the slide lists all InfiniBand port metrics. The metric values shown in the slide correspond to the amount of data being transmitted and received. For this metric, there is a metric observation for each InfiniBand port.

Quiz

Extensive retraining is required for existing Oracle database administrators to monitor and maintain Oracle databases on Exadata.

a. True
b. False

Answer: b

Summary

In this lesson, you should have learned how to describe the monitoring recommendations for Exadata database servers.

Practice 16 Overview: Oracle Database Monitoring

In this practice, you examine the Exadata-specific database monitoring capabilities provided by Enterprise Manager.

17

Monitoring the InfiniBand Network

ORACLE

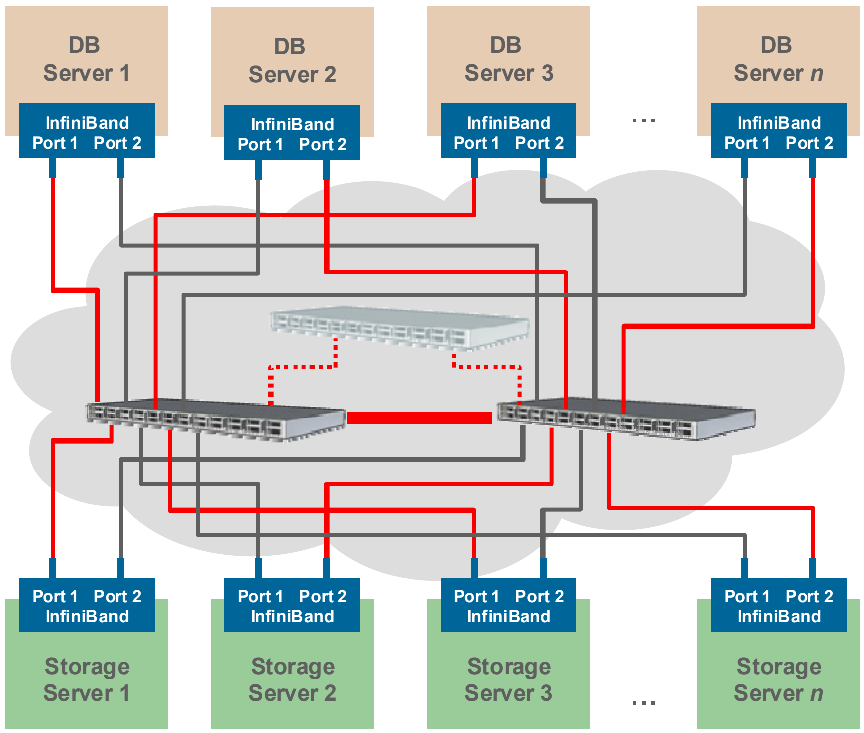
Objectives

After completing this lesson, you should be able to:

- Describe the InfiniBand network monitoring capabilities provided by Enterprise Manager Cloud Control
- Explain how to monitor the Exadata InfiniBand network without Enterprise Manager

ORACLE

InfiniBand Network Monitoring: Overview



The diagram in the slide illustrates the InfiniBand communications network contained in Exadata.

Each Exadata system contains at least two InfiniBand switches, which connect the database servers and storage servers. These are called leaf switches.

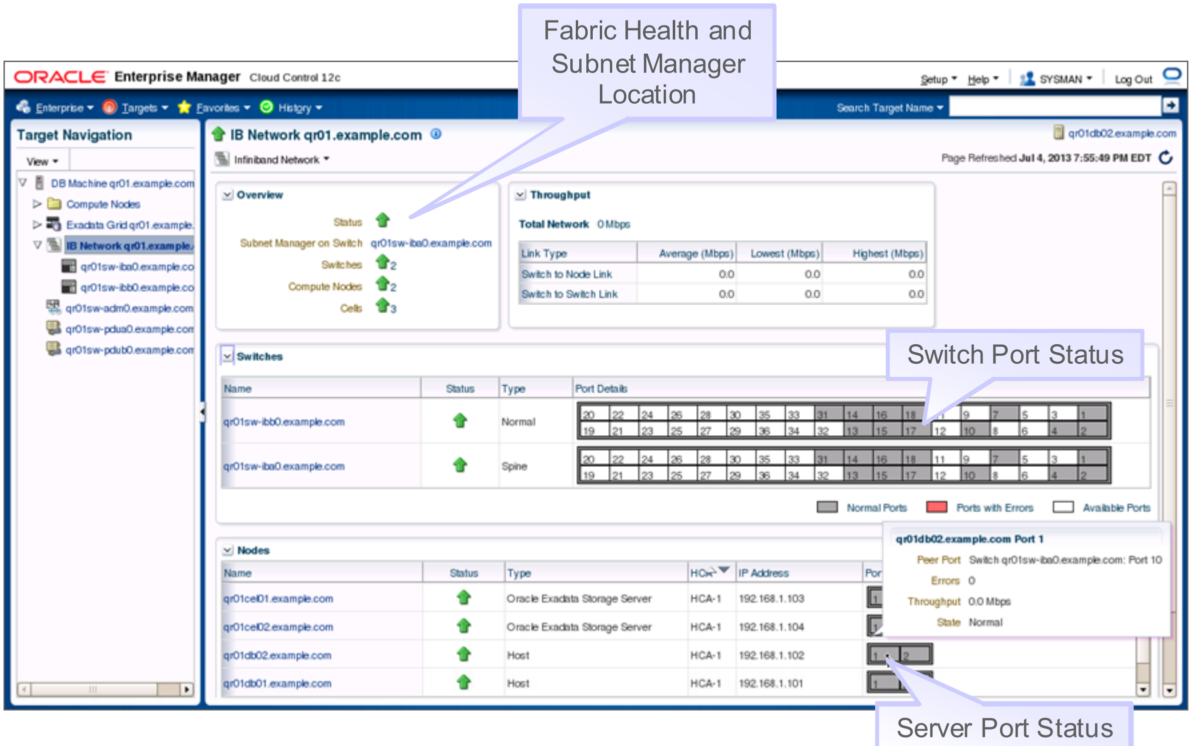
A third switch, called the spine switch, connects to both leaf switches in some older Half Rack and Full Rack Database Machines models. The spine switch facilitates the connection of multiple racks to form a single, larger Exadata environment. In newer models, that do not contain a spine switch, a spine switch can be added when required.

Each server contains at least one pair of InfiniBand ports, and all the ports are connected to the InfiniBand switches. This arrangement delivers high availability, because the system can tolerate a port failure. The port connections are spread across both leaf switches for load balancing purposes. In addition, the leaf switches within a rack are connected to each other. The result is a Fat-Tree switched fabric network topology.

Monitoring of the InfiniBand network is divided into three main areas:

- InfiniBand switch monitoring
- InfiniBand port monitoring
- Monitoring of the InfiniBand fabric

InfiniBand Network Monitoring with Enterprise Manager Cloud Control



Enterprise Manager Cloud Control provides a comprehensive monitoring capability for the Exadata InfiniBand network. Using Enterprise Manager Cloud Control, administrators can quickly and easily check the status of all the key InfiniBand network components (including switches, server ports, and the overall fabric status) on the InfiniBand Network home page. A sample screenshot is shown in the slide.

Enterprise Manager Cloud Control is equipped with a default set of metrics and threshold settings that is sufficient for general monitoring of the Exadata InfiniBand network. Thresholds can be set or altered to handle site-specific requirements.

If an error is detected, the associated switch or server ports are visually highlighted on the InfiniBand Network home page and an alert (incident) is raised.

Enterprise Manager Cloud Control also provides InfiniBand network performance monitoring capabilities, so administrators can monitor overall network performance and isolate underperforming links. Administrators can also use Enterprise Manager to:

- Enable and disable individual switch and server ports
- Reset performance and error counters
- Switch locator LEDs on and off to aid in physically locating different network ports
- Set up SNMP subscriptions on the InfiniBand switches

Monitoring the InfiniBand Switches

InfiniBand switch monitoring checks for failed switch hardware and sensors that exceed preset thresholds.

- Example of commands and expected normal output:

```
# ssh root@dm01sw-ibb0
root@dm01sw-ibb0's password:
[root@dm01sw-ibb0 ~]# showunhealthy
OK - No unhealthy sensors
[root@dm01sw-ibb0 ~]# checkpower
PSU 0 present OK
PSU 1 present OK
All PSUs OK
```

- Checks should be performed every couple of minutes.
- Use `env_test` to gather more information, if required.

ORACLE

For users without Enterprise Manager Cloud Control, InfiniBand monitoring can be achieved using various commands and utilities. The remainder of this lesson focuses on these.

Monitoring of the InfiniBand switches provided with Exadata involves checking for failed hardware components and sensors that exceed preset thresholds on the switch.

Manual monitoring can be performed by logging in to the switch as the `root` user and running the `showunhealthy` and `checkpower` commands. An example is shown in the slide. It is recommended that the checks be run every couple of minutes. In practical terms, this means creating a script to perform the check and using an automated job control system to run it repeatedly.

Note that `showunhealthy` will indicate no unhealthy sensors even if a power supply is offline, so they must be checked separately.

If `showunhealthy` or `checkpower` indicate any issues, run the `env_test` command to get detailed status information about all switch sensors.

Monitoring the InfiniBand Switch Ports

Check for switch port errors with the following command:

```
# ibqueryerrors.pl -s RcvSwRelayErrors,RcvRemotePhysErrors,XmtDiscards,
XmtConstraintErrors,RcvConstraintErrors,ExcBufOverrunErrors,VL15Dropped
```

- Compare output with previous results:
  - `SymbolErrors`, `RcvErrors`, or `LinkIntegrityErrors` should not increase without `LinkDowned` increasing.
- Checks can be executed from any database server or InfiniBand switch.
- Checks should be performed every couple of minutes.

InfiniBand port monitoring is required to check the health of InfiniBand network ports and interfaces. The recommended manual procedure is to execute the `ibqueryerrors.pl` command, as shown in the slide.

A single invocation of the `ibqueryerrors.pl` command shown in the slide can be used to report switch port errors across all the InfiniBand switches. The output of the command should be compared with previously gathered output. When all is functioning correctly, the counts for `SymbolErrors`, `RcvErrors`, or `LinkIntegrityErrors` should not increase without `LinkDowned` increasing. The check can be executed from any database server or InfiniBand switch. It is recommended that the check be performed every couple of minutes.

Monitoring the InfiniBand Ports on Exadata Servers

- InfiniBand port monitoring is automatically performed on Exadata servers.
- Manually monitor server ports with the `LIST IBPORT` command:

```
CellCLI> LIST IBPORT DETAIL
 name:                HCA-1:1
 dataRate:            "40 Gbps"
 hcaFWVersion:        2.11.1280
 id:               0x0010e000014908b9
 lid:                 18
 linkDowned:          0
 linkIntegrityErrs:   0
 linkRecovers:        0
 physLinkState:       LinkUp
 portNumber:          1
```

```
rcvConstraintErrs: 0
rcvData:           118626786672
rcvErrs:           0
rcvRemotePhysErrs: 0
status:            Active
symbolErrs:        0
vl15Dropped:       2
xmtConstraintErrs: 0
xmtData:           101435188674
xmtDiscards:       0
```


InfiniBand port monitoring on storage servers and database servers equipped with Exadata software release 12.1.2.1.0 or later, is performed by the Management Server (MS). No additional InfiniBand monitoring is required. If an InfiniBand port is not functioning correctly, the MS creates an alert, and delivers the alert to Enterprise Manager. Alert messages typically contain the corrective action to perform to resolve the alert.

You can manually monitor InfiniBand ports on Exadata servers by using the `LIST IBPORT` command. Check that every port shows the following output: `dataRate: "40 Gbps"`, `physLinkState: LinkUp`, `status: Active`. Also, the error counters (attributes with names that end in `Errs`) should not increase while the interface is operational. The slide shows some example output. Note that you can reset the InfiniBand port counters by using the `ALTER IBPORT ... RESET COUNTERS` command.

For database servers without MS, you can use the following commands and checks:

- **`ibstatus`:** Check that every port shows the following output: `state: 4: ACTIVE`, `phys state 5: LinkUp`, `rate: 40 Gb/sec (4X QDR)`.
- **`perfquery`:** `SymbolErrors` or `RcvErrors` or `LinkIntegrityErrors` should not increase without `LinkDowned` increasing.
- **`ifconfig`:** Check that the InfiniBand interfaces are `UP`.
- **`ping` and `rds-ping`:** Check for connectivity to all database servers and storage servers over the InfiniBand network.

Monitoring the InfiniBand Fabric: Subnet Manager Master Location

The following checks are recommended once per day:

- Check that the InfiniBand Subnet Manager (SM) master is located on one of the InfiniBand switches.
  - Example:

```
# sminfo
sminfo: sm lid 1 sm guid 0x21283a8516a0a0, activity count 933330 priority 5
state 3 SMINFO_MASTER
# ibswitches
Switch : 0x0021283a8516a0a0 ports 36 "Sun DCS 36 QDR switch dm01sw-ibs0"
enhanced port 0 lid 1 lmc 0
Switch : 0x0021283a8983a0a0 ports 36 "Sun DCS 36 QDR switch dm01sw-iba0"
enhanced port 0 lid 4 lmc 0
Switch : 0x0021283a89bda0a0 ports 36 "Sun DCS 36 QDR switch dm01sw-ibb0"
enhanced port 0 lid 3 lmc 0
```

- For networks containing spine switches, check that the SM master is running on a spine switch.

ORACLE

Everyday, check that the Subnet Manager (SM) master is located on one of the InfiniBand switches. From one of the database servers or InfiniBand switches, execute the `sminfo` and `ibswitches` commands as shown in the slide. Determine the host name for the SM master location by matching the `guid` reported by `sminfo` with the output from `ibswitches`. Alternatively, the `getmaster` command can be used. However, note that `getmaster` is available only on the InfiniBand switches, not the database servers.

In an Exadata environment that contains a spine switch, SM master should run on the spine switch. A spine switch is a switch that only has other switches connected to it. The following command identifies the spine switches:

```
ibnetdiscover -p | awk '/^SW +[0-9]+ +[0-9]+ +0x[0-9\
a-e]+ +[0-9]+x .DR - [SW|CA].*/ {
  if (spine[$4]=="") spine[$4]="yes"
  if ($8 == "CA") spine[$4]="no"
} END {
  for (val in spine)
    if (spine[val]=="yes")
      print val }'
```

Alternatively, current versions of the `verify-topology` command (discussed later in this lesson) can also be used to identify spine switches.

Monitoring the InfiniBand Fabric: Network Topology and Link Status

- Use `verify-topology` to check the InfiniBand fabric.
  - Example of expected output:

```
# /opt/oracle.SupportTools/ibdiagtools/verify-topology
[ DB Machine Infiniband Cabling Topology Verification Tool ]
                [Version IBD VER 2.c ]
Spine switch found: dm01sw-ibs0 (21283a8516a0a0)
Leaf switch found: dm01sw-iba0 (21283a8983a0a0)
Leaf switch found: dm01sw-ibb0 (21283a89bda0a0)
Found 2 leaf, 1 spine, 0 top spine switches
Check if all hosts have 2 CAs to different switches...............[SUCCESS]
Leaf switch check: cardinality and even distribution..............[SUCCESS]
Spine switch check: Are any Exadata nodes connected ..............[SUCCESS]
Spine switch check: Any inter spine switch links..................[SUCCESS]
Spine switch check: Any inter top-spine switch links..............[SUCCESS]
Spine switch check: Correct number of spine-leaf links............[SUCCESS]
Leaf switch check: Inter-leaf link check..........................[SUCCESS]
Leaf switch check: Correct number of leaf-spine links.............[SUCCESS]
```

- Use `iblinkinfo.pl -Rl` to monitor the link status.
  - Compare output with previously gathered output.

ORACLE

In addition to checking the SM master, it is recommended to monitor the InfiniBand network topology and the state of all the InfiniBand network links.

It is recommended to execute the `verify-topology` command from one of the Exadata database servers, once per day. `verify-topology` conducts a series of topology-level tests to verify that the expected storage network links are cabled correctly. An example of the expected output for `verify-topology` is shown in the slide.

To check the state of all the InfiniBand network links, run the following command as `root` on one of the Exadata database servers:

```
# iblinkinfo.pl -Rl
```

Compare the resulting output with previously gathered output from a time when the InfiniBand network was in a known, good state. It is recommended to run this check frequently (every few minutes).

Quiz

Enterprise Manager Cloud Control provides a comprehensive set of Exadata InfiniBand network monitoring capabilities.

a. True

b. False

Answer: a

Summary

In this lesson, you should have learned how to:

- Describe the InfiniBand network monitoring capabilities provided by Enterprise Manager Cloud Control.
- Explain how to monitor the Exadata InfiniBand network without Enterprise Manager.

ORACLE

Practice 17 Overview: InfiniBand Monitoring

In this practice, you examine the InfiniBand network monitoring capabilities provided by Enterprise Manager.

18

Monitoring Other Exadata Database Machine Components

ORACLE

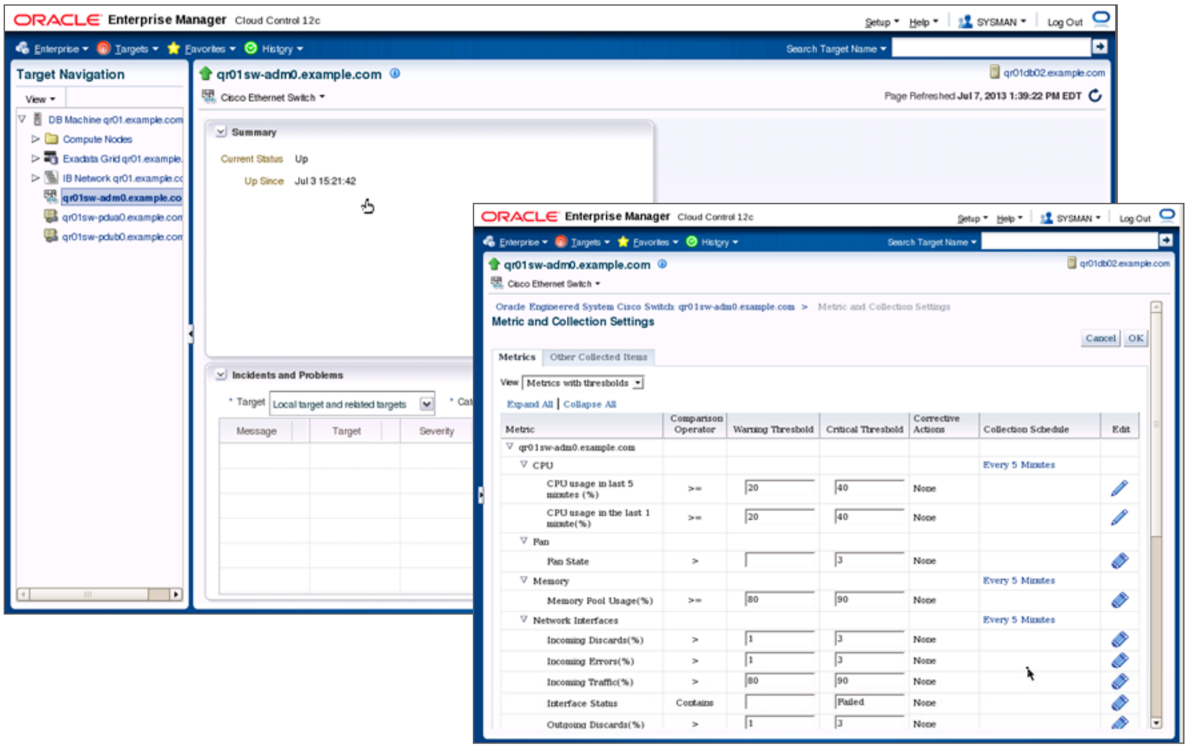
Objectives

After completing this lesson, you should be able to monitor the following additional Exadata components:

- Cisco Ethernet Switch
- Sun Power Distribution Units
- KVM Switch

ORACLE

Monitoring the Cisco Ethernet Switch



The Cisco Ethernet switch is not in the critical data path. It does not participate in connecting application clients to the database or in connecting database servers to the storage servers. However, monitoring and administrative traffic depends on the availability of the Cisco switch.

The primary goal of monitoring the Cisco switch is to identify hardware component failure and environmental conditions that can lead to switch malfunction. The switch monitors availability and sensor thresholds for its hardware components.

Enterprise Manager facilitates easy monitoring of the Cisco switch. An example of the Cisco switch home page is shown in the slide. The metrics and default thresholds provided by Enterprise Manager are sufficient for the level of monitoring that is necessary to ensure switch availability. Thresholds can be changed or set on the Metric and Collection Settings page to handle site-specific requirements. Alerts within the switch are reported using an SNMP trap, which can be received and displayed by Enterprise Manager.

Manual monitoring of the Cisco switch can be achieved using a command-line interface that is built into the Cisco switch. For more information, refer to the documentation for Cisco Catalyst 4900 Series switches, which is available at http://www.cisco.com/go/4900.

Monitoring the Power Distribution Units

- Each rack contains two PDUs.
- Monitor PDUs to:
  - Ensure continuous power supply
  - Measure power consumption
- Monitoring options:
  - Physical inspection of the PDUs
  - Remote monitoring using Enterprise Manager or other SNMP manager
    - Appropriate threshold settings inside each PDU are required to facilitate remote monitoring.

ORACLE

Every Exadata rack contains two Power Distribution Units (PDUs). The PDUs provide a redundant power supply capability to maintain high availability of the system. The primary goal of monitoring the PDUs is to ensure that a continual power supply is maintained. PDU monitoring can also measure power usage for cost management purposes.

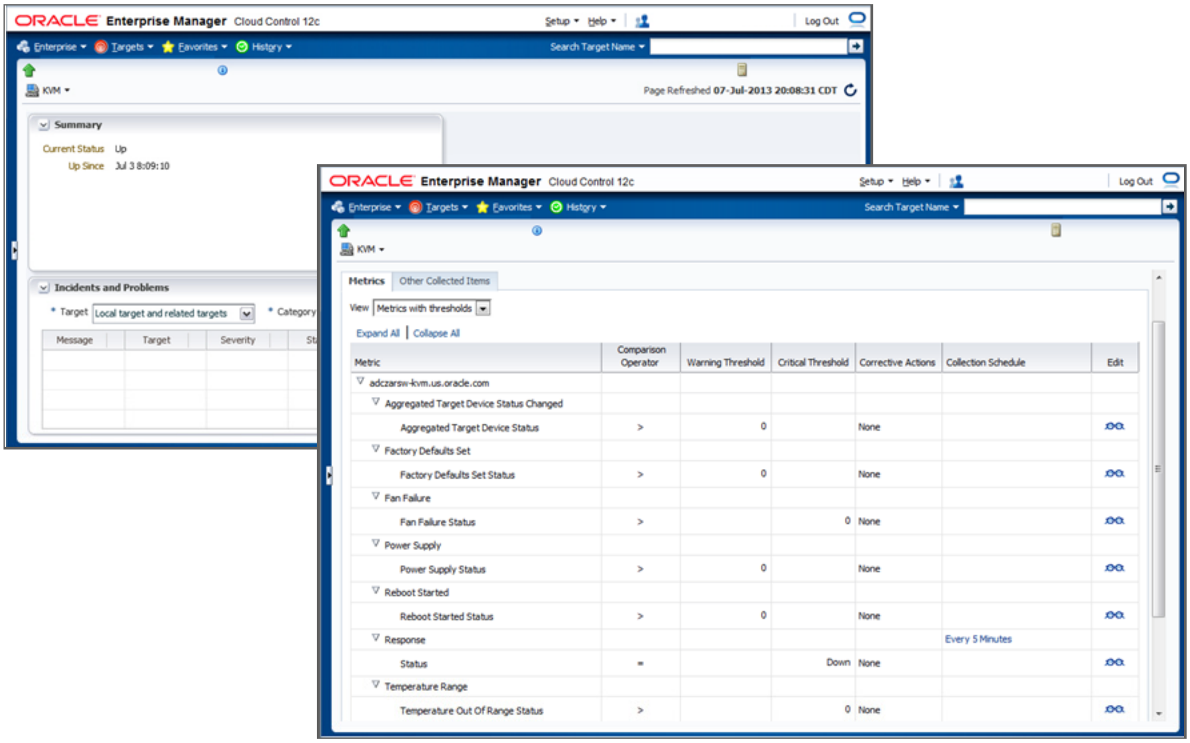
On each PDU, the PDU metering unit monitors the current being used by the equipment connected in the rack. You can monitor the power usage in person by viewing the LCD screen on the PDU or remotely from a system on the network. A picture of the PDU power consumption monitor is shown in the slide.

Enterprise Manager can also be used to monitor PDU availability and to view any alerts generated by the PDU.

For network-based PDU monitoring, with or without Enterprise manager, each PDU needs an IP address allocation and should be connected directly to your administration network, and not the Cisco Ethernet switch. Through this configuration, you can quickly determine if a system outage is due to an issue with the power supply.

Appropriate thresholds must be configured in each PDU to automatically generate alerts indicating abnormal power consumption. The specific threshold values depend on several factors including the model of the Database Machine, the size of the Database Machine, and the power supply voltage and type. The *Oracle Exadata Database Machine Installation and Configuration Guide* contains the recommended threshold values for each situation.

Monitoring the KVM Switch



Some Exadata X2 and V2 models include a KVM switch. It's primary purpose is to provide keyboard, video, and mouse (KVM) control to facilitate console access to the Database Machine database servers and Exadata Storage Servers. No KVM switch exists on Exadata X2-8, X3, or newer models. On these models, server console functionality can be accessed using ILOM.

The KVM switch is monitored to ensure availability. The KVM switch can be monitored using Enterprise Manager. An example of a KVM switch home page is shown in the slide. The metrics and default thresholds provided by the plug-in are sufficient for the level of monitoring necessary to ensure switch availability. Alerts within the switch are reported using an SNMP trap, which can be received and displayed by Enterprise Manager.

Quiz

Identify the component that is most important to monitor to maintain the overall availability of Database Machine.

a. Cisco Ethernet Switch
b. Sun Power Distribution Units
c. KVM Switch

Answer: b

Summary

In this lesson, you should have learned how to monitor the following additional Exadata components:

- Cisco Ethernet Switch
- Sun Power Distribution Units
- KVM Switch

19

Other Useful Exadata Monitoring Tools

ORACLE

Objectives

After completing this lesson, you should be able to describe the following tools:

- Exachk
- ExaWatcher
- TFA Collector
- ADRCI
- Imageinfo and Imagehistory
- ILOM

Exachk: Overview

- Exachk is a utility that:
  - Collects data regarding Exadata component versions and best practices
  - Is used by administrators to check their environment against supported version levels and best practices
  - Is preinstalled on new Exadata Database Machines
  - Is available from My Oracle Support note 1070954.1
  - Should be executed periodically as a regular part of Exadata monitoring and maintenance
  - Does not alter any Exadata configuration settings
  - Is lightweight and has minimal impact on the system
- Exachk is NOT:
  - A continuous monitoring tool
  - A replacement for Enterprise Manager

ORACLE

Exachk collects data regarding key software, hardware, and firmware versions, along with information about Exadata-specific configuration best practices.

The output assists Exadata administrators to periodically review and cross-reference the current data for key Exadata components against supported version levels and recommended best practices.

Exachk is preinstalled on new Exadata systems. The latest updated version of Exachk is available for download from My Oracle Support note 1070954.1. It can be executed as desired and should be executed regularly as part of the maintenance program for Exadata.

Exachk is not a continuous monitoring utility and should not be considered as a replacement for other monitoring or alerting tools, such as Enterprise Manager or ILOM.

It is non-intrusive and does not change anything on Exadata apart from writing output files and some small temporary files, which are deleted after they are used. It also offers to configure SSH user equivalence if it is not configured.

Exachk is lightweight. Therefore, the impact to the target machine is minimal. However, it is thorough, and consequently takes some time to execute. In general, Exachk takes a few minutes for each database server in your environment. However, the amount of time taken can vary depending on numerous factors, such as the number of nodes in the cluster, the current CPU load, the current network latency, and so on.

Running Exachk

- Run exachk:
  - Typically located under `/opt/oracle.SupportTools/exachk`
  - Beginning with exachk version 12.1.0.2.2, it is recommended to execute exachk as the `root` user.
    - Older versions can be run as the `oracle` or `grid` user.
  - Follow the prompts, and read and understand all the messages.
- Review the report:
  - Review the summary to identify areas for further investigation.
  - Review the details for recommendations and further information.

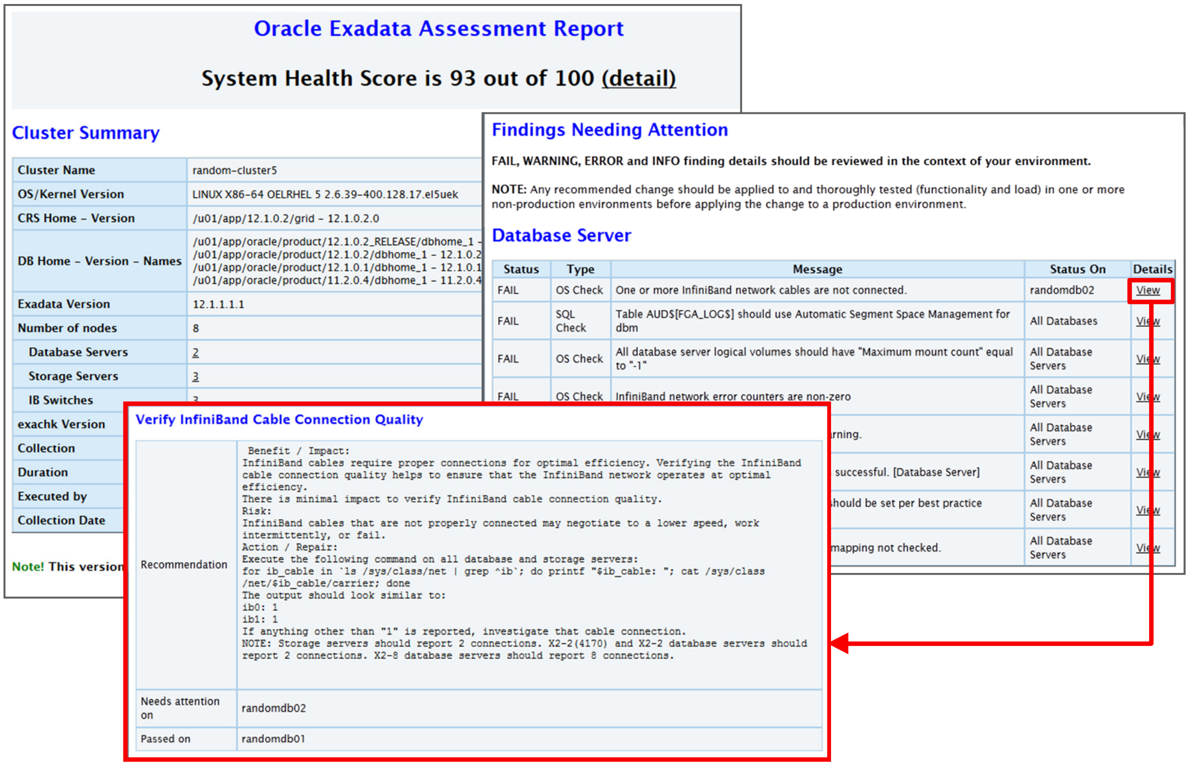
The exachk executable does not require any mandatory command-line options because the tool automatically scans Exadata to gather most of the required information. Additional information is gathered using prompts at the beginning of the exachk session. After the administrator input is gathered, exachk can be left to run unattended.

The following usage considerations should also be noted:

- Beginning with exachk version 12.1.0.2.2, it is recommended to execute exachk as the `root` OS user. The exachk process running under the `root` user id uses `su` to execute commands as the Oracle Database software owner (typically, the `oracle` OS user) or the Grid Infrastructure software owner (typically, the `grid` OS user) as required.
- In a non-role-seperate environment, older versions of exachk can be run as the `oracle` user. In a role-separated environment, older version of exachk must be installed and run separately by the `oracle` and the `grid` users.
- To maximize the number of checks that are performed, execute exachk when the Grid Infrastructure services and at least one database are up and running.
- Although exachk is a minimal impact tool, it is best practice to execute it during times of least load on the system.

- To avoid possible problems associated with running the tool from terminal sessions on a network-attached computer, run the tool using VNC. Hence, if there is a network interruption, the tool will continue to execute to completion.
- If the execution of the tool fails for some reason, it can be rerun from the beginning. Exachk does not resume from the point of failure.

Exachk Output



ORACLE

Exachk produces an HTML report of findings, with the most important exceptions listed first by component. The report is contained in a date-stamped zip file that is produced during each exachk run. Typically, administrators will transfer the exachk output ZIP file to a desktop computer and expand it there to review the exachk findings.

The exachk HTML report contains a summary of findings, along with relevant details. The detailed entries typically describe the issues in detail and provide recommendations for remediation. A link to documentation or a link to a note in My Oracle Support may also be included to provide further information, especially if the recommendations are not straightforward.

The slide shows an example of the output from the exachk HTML report.

Exachk Daemon

Exachk version 2.2.2 introduced a daemon process to permit non-interactive execution at regular intervals.

- Configuration example:

```
# ./exachk -set "AUTORUN_INTERVAL=1d;AUTORUN_FLAGS= -o -v
;NOTIFICATION_EMAIL=name@company.com;PASSWORD_CHECK_INTERVAL=1"
```

- Checking the configuration:

```
# ./exachk -get all
```

- Starting and stopping the daemon:

```
# ./exachk -d [ start | stop ]
```

- Checking the daemon status and the next scheduled run:

```
# ./exachk -d [ status | nextautorun ]
```

ORACLE

Exachk version 2.2.2 introduced a daemon process to permit non-interactive execution at regular intervals.

Before running exachk in daemon mode, administrators must set key parameters by running exachk with the `-set` command-line option. Multiple parameters, delimited by semicolons, can be set at once. (An example is displayed in the slide.) Run `exachk -h` for a list of the available parameters.

After the daemon configuration is completed, you can start, stop, and check the daemon status by using the commands shown in the slide. Starting the daemon also executes the first scheduled exachk run.

Note the following regarding the exachk daemon:

- After the daemon is started, you can initiate an exachk run manually on demand by running the exachk script.
- The daemon can be configured to automatically restart if the server is rebooted.
- The daemon must be restarted if the Exadata configuration changes (For example, servers are added or removed).

ExaWatcher: Overview

- ExaWatcher collects operating system and network metrics on all Exadata servers to aid with diagnosis of performance issues.
- ExaWatcher gathers data on a regular basis, invoking utilities such as `vmstat`, `netstat`, and `iostat`.
- ExaWatcher runs automatically with the following default settings:
  - Installation location: `/opt/oracle.ExaWatcher`
  - Most OS commands are executed every five seconds.
    - A few long-running commands have a longer interval.
  - Archive location: `/opt/oracle.ExaWatcher/archive`
    - 3 GB on database servers, 600 MB on storage servers
- You can customize the commands that are executed, collection intervals, the archive size, and more.
- Use the `GetExaWatcherResults.sh` script to search the archive and write a set of observations to a specific location.
- See My Oracle Support note 1617454.1 for more details.

ExaWatcher is a collection of perl and shell scripts designed to collect and archive operating system and network metrics to aid in diagnosing performance issues. ExaWatcher replaces OSWatcher in Exadata software versions 11.2.3.3 and later. ExaWatcher operates as a set of background processes on the server and gathers OS data on a regular basis, invoking utilities such as `vmstat`, `netstat`, and `iostat`.

By default, metric snapshots for most commands are recorded every five seconds. However, some long-running commands have a longer interval between executions. For example, the list of files opened by current processes, as determined by the `lsof` command, has a minimum collection internal of two minutes.

The default space allocation for the ExaWatcher archive is 3 GB on each database server and 600 MB on each storage server. Typically, the default allocation is large enough to store approximately two weeks of observations on each storage server, and approximately 10 weeks of observations on each database server. ExaWatcher periodically checks the amount of space consumed by the archive and deletes old observations to create space, if required.

ExaWatcher is installed under `/opt/oracle.ExaWatcher`. Metric observations are maintained in the `archive` subdirectory.

See My Oracle Support note 1617454.1 for more information about ExaWatcher, including how you can use the `GetExaWatcherResults.sh` script to search the archive and write a set of observations to a specific location, and how you can customize ExaWatcher.

TFA Collector: Overview

- TFA Collector collects diagnostic data:
  - From Oracle and OS logs, traces, and dumps
  - Across the entire Exadata Database Machine
- TFA Collector generates diagnostic data packages:
  - Automatically when a critical event occurs
  - On-demand for your specified time interval
- Features:
  - A single command performs clusterwide diagnostic collection for all components
  - Diagnostic files are trimmed around the incident time
  - Collected diagnostics are consolidated on a single node
  - Integrated with Exachk and ExaWatcher
- Refer to My Oracle Support note 1513912.2 for more details.

Trace File Analyzer (TFA) Collector is a utility that simplifies diagnostic data collection on clustered Oracle systems. TFA Collector collects and packages diagnostic data from numerous Oracle and Operating System log files, trace files, and dump files.

TFA Collector maintains a historical repository of diagnostic data, and provides commands that allow you to easily generate a diagnostic package for a specific time window, even trimming diagnostic files to your specifications. TFA Collector automatically generates a diagnostic package when a critical event occurs, to ensure that diagnostic data is not lost or overwritten. TFA Collector implements flood control to ensure that it does not flood system resources while collecting diagnostic data, and so you can control TFA Collector from a central location. TFA Collector is integrated with Exadata, being able to collect diagnostic data from Exadata Storage Servers, and also collect information generated by Exachk and ExaWatcher.

The TFA Collector binaries are included as part of the Oracle Database and Oracle Grid Infrastructure distributions since versions 11.2.0.4 and 12.1.0.1. On Exadata, TFA Collector is automatically installed and configured on the database servers as part of the Grid Infrastructure installation on new Exadata systems. However, you must perform a one-off manual configuration, so TFA Collector can perform diagnostic collections on the Exadata Storage Servers. For systems without TFA Collector, you can download and install it.

The TFA Collector software and user guide is available at My Oracle Support note 1513912.2.

Running TFA Collector on Exadata

- TFA Collector starts and stops automatically on reboot
- Binaries are typically located under the Grid Infrastructure home directory
- Repositories are typically located under `/u01/app/oracle/tfa/repository`
- TFA Collector is controlled using `tfactl`:
  - Commands must be run by the `root` OS user for access to full functionality and diagnostics

    ```
    # tfactl <command> [options]
    ```

  - Key commands:
    - `diagcollect` - Perform on-demand diagnostic collection
    - `analyze` - List event summary and search for strings in log files
    - `cell` - Configure or modify Exadata Storage Cell features
    - `print` - Print requested details
    - `start`, `stop` - Manually start or stop TFA Collector

ORACLE

TFA Collector consists of a daemon, and a Command Line Interface (CLI). After configuration, the TFA daemon (`TFAMain`) automatically runs at all times by default on each Exadata database server. The TFA daemon does not run on the Exadata Storage Servers. Instead, the daemons on the database servers connect to the Exadata Storage Servers by using SSH to collect diagnostic information from the cells.

The TFA Collector binaries are typically located under the Grid Infrastructure home directory on newly installed Exadata systems. If you add TFA Collector to an existing system, you can specify the installation location. TFA Collector maintains a repository of diagnostic information, which is typically located under `/u01/app/oracle/tfa/repository` on each server running `TFAMain`.

The TFA Collector CLI is `tfactl`. Because TFA Collector accesses privileged OS diagnostic data, most `tfactl` commands must be run as the `root` OS user. The slide shows the general format of the `tfactl` command. If you execute `tfactl` without specifying a command, you enter a command processor where you can execute a series of commands.

Some of the key commands are outlined in the slide. Some of the commands have numerous options. For example, the `diagcollect` command specifies various parameters for your diagnostic collection, including the time period, software components, databases, and servers that are involved. The full command reference is contained in the TFA Collector User Guide, which is available from My Oracle Support note 1513912.2.

Using ADRCI on Exadata Storage Servers

- The Exadata Storage Server software uses the Automatic Diagnostic Repository (ADR) structure to manage diagnostic data.
- Using ADRCI on an Exadata Storage Server is essentially the same as using it on a database server.
- The default ADR home location is `/opt/oracle/cell<version>/log/diag/asm/cell/<hostname>`
- Refer to the demonstration: Using ADRCI on an Exadata Storage Cell. http://apex.oracle.com/pls/apex/f?p=44785:24:1974365757445289::NO:24:P24_CONTENT_ID,P24_PREV_PAGE:5521,29

ORACLE

The Automatic Diagnostic Repository (ADR) is a file-based repository for diagnostic data, such as traces, dumps, alert logs, health monitor reports, and more. The Exadata Storage Server software uses an ADR structure that is essentially the same as the ADR used by the Oracle Database software. Like the Oracle Database ADR, the cell ADR can be managed by using the ADR command interpreter utility (ADRCI).

Using ADRCI on an Exadata Storage Server is essentially the same as using it on a database server. The main difference is the location of the ADR. On Exadata Storage Servers, `/opt/oracle/cell<cell_version>/log/diag/asm/cell/<hostname>` is the default ADR home location, where `<cell_version>` is a string denoting the cell software version and `<hostname>` is the name of the cell host. Note that each Exadata cell has a symbolic link at `/opt/oracle/cell` that points to the current software version directory (`/opt/oracle/cell<cell_version>`).

An example of how to use ADRCI to collect trace files for a particular incident on an Exadata Storage Server is provided in the demonstration titled, *Using ADRCI on an Exadata Storage Cell*.

imageinfo: Overview

```
[root@exa1cel01 ~]# imageinfo

Kernel version: 2.6.39-400.248.3.el6uek.x86_64 #1 SMP Wed Mar 11 18:04:34 PDT 2015
x86_64
Cell version: OSS_12.1.2.1.1_LINUX.X64_150316.2
Cell rpm version: cell-12.1.2.1.1_LINUX.X64_150316.2-1.x86_64

Active image version: 12.1.2.1.1.150316.2
Active image activated: 2015-09-23 03:07:28 -0700
Active image status: success
Active system partition on device: /dev/md5
Active software partition on device: /dev/md7

Cell boot usb partition: /dev/sdm1
Cell boot usb version: 12.1.2.1.1.150316.2

Inactive image version: undefined
Rollback to the inactive partitions: Impossible
```

```
[root@exa1db01 ~]# imageinfo

Kernel version: 2.6.39-400.248.3.el5uek #1 SMP Wed Mar 11 18:14:36 PDT 2015 x86_64
Image version: 12.1.2.1.1.150316.2
Image activated: 2015-09-23 03:11:54 -0700
Image status: success
System partition on device: /dev/mapper/VGExaDb-LVDbSys3
```

ORACLE

Software is initially installed on Exadata storage and database servers using an Exadata-specific imaging process. Exadata Storage Server patch bundles are also provided as software images.

Two utilities are provided to monitor the Exadata images:

- **`imageinfo`:** Displays information relating to the images currently installed on a server
- **`imagehistory`:** Displays historical information about all the images installed on the server

The slide shows examples of the output displayed by `imageinfo` on a storage server and on a database server.

There is significantly less output associated with a database server because database servers have no concept of active and inactive images, and they do not contain a cell boot USB device.

The behavior and output of `imageinfo` can be modified using a series of optional parameters. Execute `imageinfo -h` to obtain a list of the available options.

imagehistory: Overview

```
[root@exa1cel01 ~]# imagehistory
Version                              : 11.2.2.2.0.101206.2
Image activation date                : 2011-01-12 14:56:44 -0800
Imaging mode                         : fresh
Imaging status                       : success

Version                              : 11.2.2.2.1.110131
Image activation date                : 2011-02-23 18:30:44 -0800
Imaging mode                         : out of partition upgrade
Upgrade logs                         : /var/log/cellos/patch/rollback_20
1102242002_11.2.2.2.0.101206.2_11.2.2.2.1.110131.tar.gz
Imaging status                       : success

Version                              : 11.2.2.2.0.101206.2
Image activation date                : 2011-02-24 12:22:44 -0800
Imaging mode                         : out of partition rollback
Imaging status                       : success
```

ORACLE

`imagehistory` displays historical information about all the images installed on the server.

The slide shows an example of the output displayed by `imagehistory`. In this example, the command shows information associated with a storage server. From the output, you can see that the storage server was initially imaged with the storage server software version 11.2.2.2.0. Then it was patched to version 11.2.2.2.1. And finally, the patch was rolled back leaving the server with the initial software version.

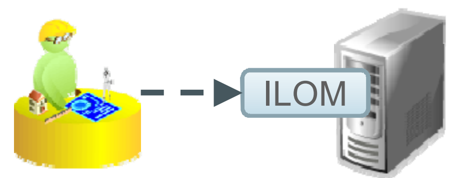
Typically, the output associated with a database server would only relate to the initially installed image because database servers are not patched using imaging techniques.

`imagehistory` has only three optional parameters:

- **`-all`**: Shows all the available image attributes
- **`-help`**: Shows basic command using instructions
- **`-version <version>`**: Outputs information related to the specified version only

Integrated Lights Out Manager (ILOM): Overview

- Is available on all Oracle hardware
- Is accessed on Exadata from the management network
- Enables system administrators to remotely:
  - Power the device on or off
  - Determine the current hardware, firmware, and network configurations of the device
  - Obtain information about hardware errors and faults
  - Configure alerts for system events
  - Upgrade device firmware
  - Browse device logs



ORACLE

The system monitoring features in ILOM easily determine the health of the system and detect errors at a glance, when they occur.

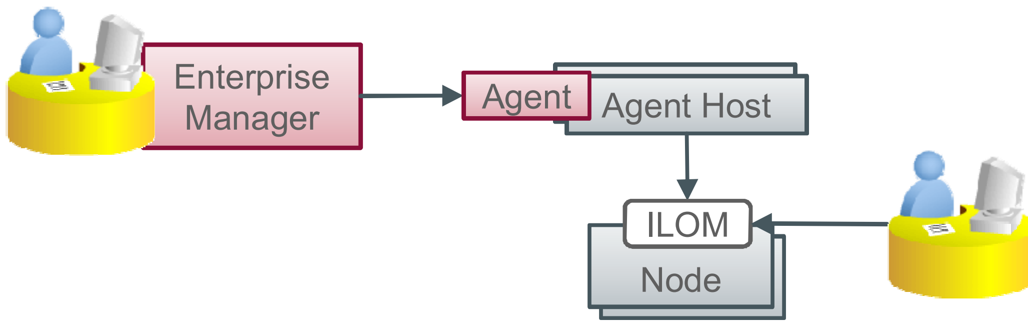
Exadata servers and InfiniBand switches are equipped with several sensors that measure voltages, temperatures, fan speeds, and other attributes of the system. Each sensor is associated with attributes such as sensor type, sensor class, threshold values and the actual sensor value. Additionally, if an alert rule is associated with a threshold value, ILOM automatically generates an alert message when the threshold is crossed.

ILOM receives error telemetry about the error events that occur within the major system components on the host (CPU, memory, and I/O hub) and the environmental subsystem within the chassis (such as fans, power supplies, and temperature). The components and conditions are constantly analyzed and fault events are captured in the ILOM event log.

On Exadata, you will typically establish communication with ILOM by using the dedicated ILOM network port, which is connected to the management network.

Using ILOM

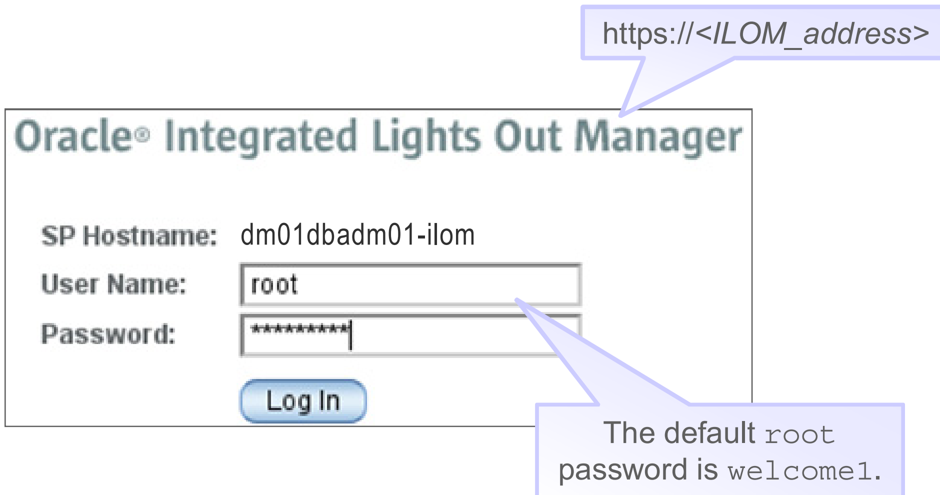
- Monitor and trigger alerts for Exadata server hardware with Enterprise Manager Cloud Control.
- Directly access a device's browser or command-line ILOM interface:
  – If Cloud Control is not available
  – To power the device on/off
  – To modify the device's ILOM configuration



Enterprise Manager Cloud Control communicates with the ILOM interface on Exadata servers and switches in order to monitor the hardware. ILOM monitors the health of hardware components like power supplies, CPUs, RAM, disks, and PCI cards. Enterprise Manager communicates with the hardware by using agents that are configured on a set of the Exadata database servers.

Alternatively, administrators can directly access the ILOM by using a Web interface or by using a command-line interface (CLI).

Accessing the Browser Interface



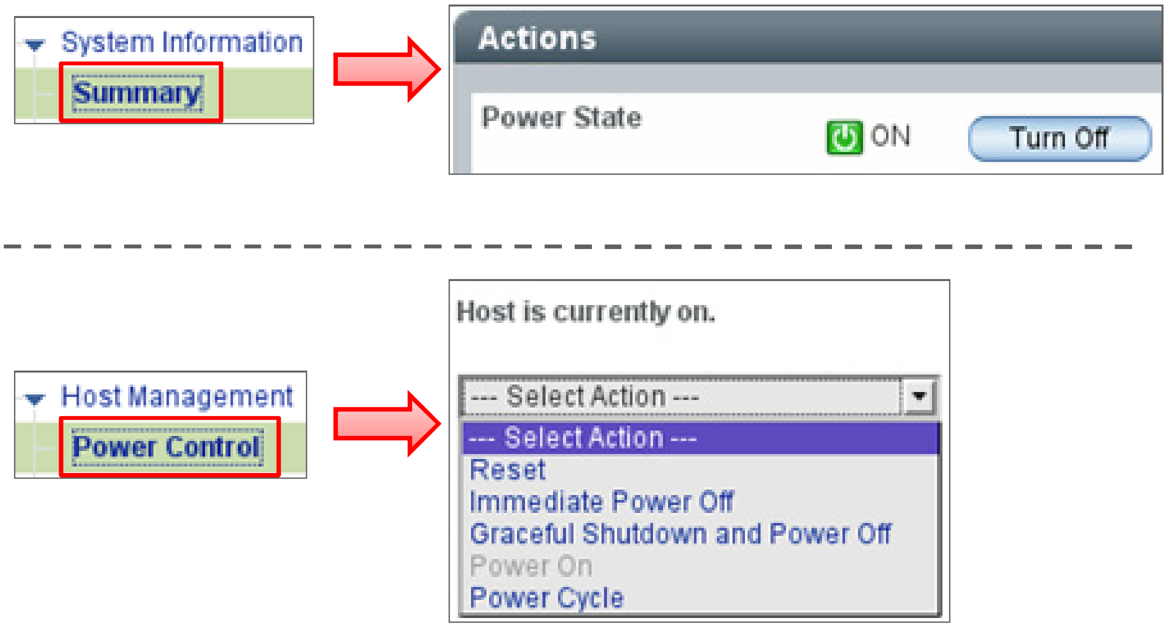
ORACLE

Open a web browser and connect to the ILOM web interface by specifying the management controller's network address in the URL. The default ILOM host names in Exadata are of the form `<host>-ilom`. For example, `dm01dbadm01-ilom`. The default credentials are the username `root` and the password `welcome1`.

ILOM supports a maximum of 10 active user sessions, including serial, Secure Shell (SSH), and web interface sessions per service processor (SP).

The ILOM web interface uses a model of hierarchical tabbed pages that you select by clicking the tab name. When a page is displayed, you can provide information, set parameters, or access other subtabs. For some pages, initiating a task might spawn an additional window, which accepts additional parameters.

Powering the Device On or Off



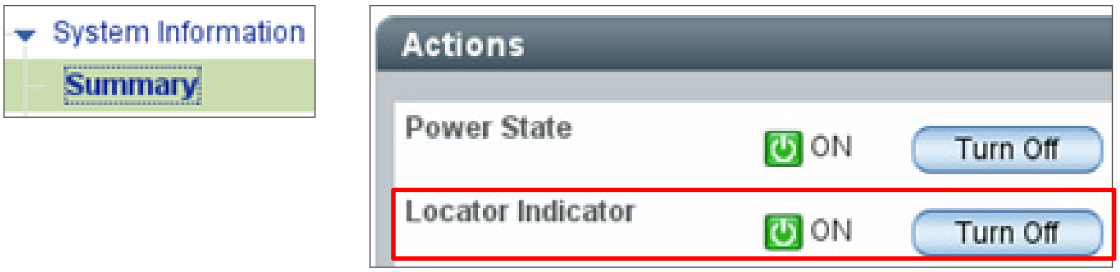
There are two pages in the browser interface that power on and power off the host hardware. The home page (System Information > Summary) includes an Actions area. Click Turn Off to power off the device and Turn On to power it back on. You can also use the Locator Indicator action to help identify a device within the Exadata rack.

The Turn Off button performs a graceful shutdown of the operating system before powering off the host server.

Alternatively, click Host Management > Power Control. Select one of the following actions and click Save:

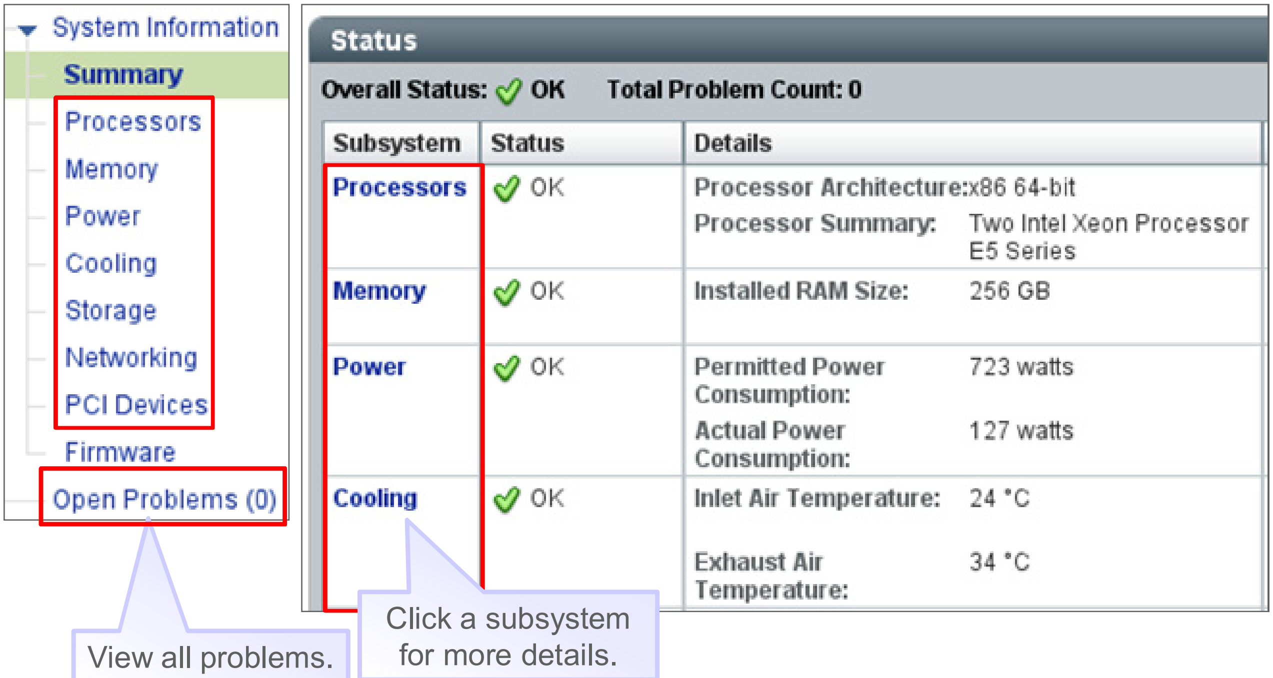
- **Reset:** This option cycles power on the host server, while keeping power applied to system components (such as disk drives).
- **Immediate Power Off:** This option immediately turns off the power on the host server.
- **Graceful Shutdown and Power Off:** This option shuts down the OS gracefully before powering off the host server.
- **Power On:** This option turns on full power to a host server that is running on standby power.
- **Power Cycle:** This option turns off system power to all system components, and then applies full power to all system components.

Locating the Device



To locate a specific Exadata component in a rack, you can use either the Oracle ILOM CLI or web interface to illuminate the Locator LED that is typically located on both the front and back panels of the device.

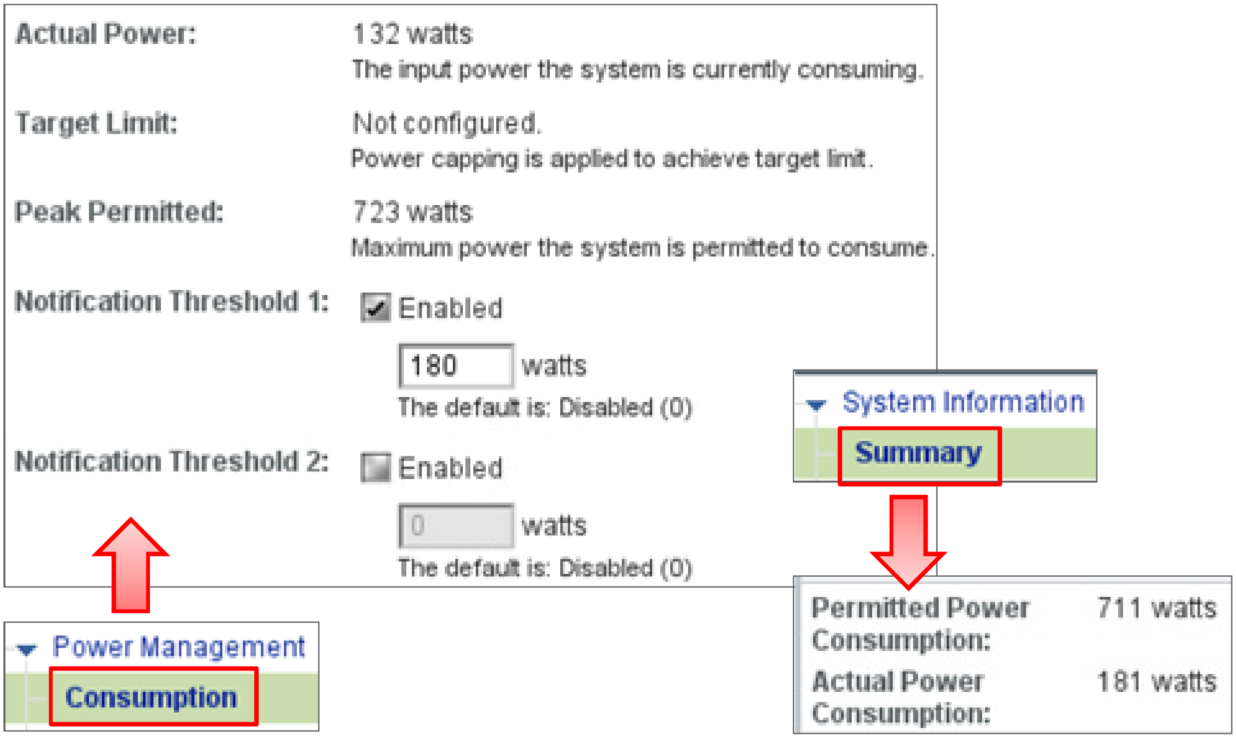
Viewing Hardware Status and Specifications



ILOM regularly polls the sensors in the system and reports any events it encounters about sensor state changes or sensor threshold crossings to the ILOM event log. Additionally, if an alert rule is enabled, ILOM automatically generates an alert message to the alert destination that you define when the associated threshold is crossed.

From the web interface Summary page, you can collect system-specific information that describes the managed device, determine the health of the managed device, and view problems on a managed device. The overall health status and total problem count appear at the top of the table. To view additional information about a subcomponent category that is reported in the Status table, click the corresponding link in the Subsystem column.

Monitoring Power Consumption



ORACLE

Monitor the current power consumption and configure power notifications by using the Power Management > Consumption page.

Actual Power represents the consumed power for the server.

Peak Permitted power represents the maximum input power that the server guarantees it will consume at any instant. Hardware maximum power identifies the maximum input power that a system is capable of consuming at any instant, given the hardware configuration of the system. Therefore, the hardware configuration maximum power is the sum of the maximum power that each processor, I/O module, memory module, fan, and so forth is capable of consuming.

By using Notification Threshold settings, you can generate two power consumption notifications when the specified power consumption value exceeds the threshold. An ILOM event is then generated, which can in turn trigger an alert.

Accessing the Command-Line Interface (CLI)

SSH to the ILOM address:

```
$ ssh root@dm01dbadm01-ilom
...
-> show /System
   Targets:
      Open_Problems (0)
      Processors
      Memory
   ...
   Properties:
      health = OK
      open_problems_count = 0
   ...
   Commands:
      cd
      reset
      set
      show
```

OR

Use the `ipmitool` command:

```
$ ssh root@dm01dbadm01
...
$ ipmitool sunoem cli
-> show /System
...
```


Use the `help` command to display help information about commands and targets. The commands that are available include `cd`, `create`, `delete`, `exit`, `help`, `load`, `reset`, `set`, `show`, `start`, `stop`, and `version`. The `ls` command performs the same function as the `show` command.

The CLI consists of a hierarchy of targets. For each target, you can view its child targets, its properties, and the commands that you can execute on the target. The CLI supports tab completion when entering target names.

CLI: Examples

Change a user's password:

```
set /SP/users/root password
Enter new password: ********
```

Monitor hardware:

```
show /System/Processors/CPUs/CPU_0
...
health = OK
part_number = 060E
model = Intel(R) Xeon(R) CPU E5-2697 v2 @ 2.70GHz
```

Monitor power consumption:

```
show /System/Power
...
health = OK
actual_power_consumption = 194 watts
```

ORACLE

The targets and properties below `/SP` are used for configuring the ILOM service processor (SP) and for viewing logs and consoles.

The targets and properties below `/System` provide hardware inventory, monitoring, and management. Hardware components such as memory and fans are named and organized in the same manner as the browser interface.

The targets and properties below `/HOST` are used for monitoring and managing the host operating system.

Quiz

Exachk is an alternative to using Enterprise Manager.

a. True

b. False

Answer: b

Quiz

Which utility should you use to provide Oracle Support with a consolidated package of information about an Exadata system?

a. Exachk
b. ExaWatcher
c. TFA Collector
d. ADRCI
e. Imageinfo
f. Imagehistory
g. ILOM

Answer: c

Quiz

Identify the tasks that you can perform with ILOM.

a. Analyze power consumption.
b. Monitor hardware sensors.
c. Create OS users and roles.
d. Repair failed hardware.

Answer: a, b

Summary

In this lesson, you should have learned how to describe the following tools:

- Exachk
- ExaWatcher
- TFA Collector
- ADRCI
- Imageinfo and Imagehistory
- ILOM

Additional Resources

- Lesson demonstrations
  - Using ADRCI on an Exadata Storage Server
    http://apex.oracle.com/pls/apex/f?p=44785:24:1974365757445289::NO:24:P24_CONTENT_ID,P24_PREV_PAGE:5521,29
- My Oracle Support notes
  - Oracle Exadata Database Machine exachk or HealthCheck
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1070954.1
  - TFA Collector - Tool for Enhanced Diagnostic Gathering
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1513912.2
  - ExaWatcher utility on Exadata database servers and storage cells
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1617454.1

ORACLE

20

Backup and Recovery

Objectives

After completing this lesson, you should be able to:

- Explain how Recovery Manager (RMAN) backups are optimized using Exadata Storage Server
- Describe the recommended approaches for disk- and tape-based backups of databases on Exadata
- List the recommended best practices for backup and recovery on Exadata

ORACLE

Backup and Recovery: Overview

- Backup and recovery of databases on Exadata:
  - Use RMAN
  - Typical strategies:
    - Disk-based backups
    - Tape-based backups
    - Hybrid strategy
- Backup and recovery of Exadata software:
  - Database server software
  - Exadata Storage Server software

This lesson focuses on the best practice recommendations for backup and recovery in conjunction with Exadata. In particular, the recommendations throughout this lesson relate to Oracle Database version 11.2.0.2 and later. The slide lists the topics that are considered throughout the lesson.

Using RMAN with Exadata

- Using RMAN on Exadata is essentially the same as using RMAN elsewhere:
  - Same concepts
  - Same commands
- Incremental backup performance is improved.
  - Block filtering is offloaded to Exadata Storage Server.
  - Fewer blocks need to be processed by RMAN.
  - Offload processing is automatic and transparent.
- Hybrid Columnar Compression can assist to further improve backup performance.
  - Reduced data size results in smaller, quicker backups.

Oracle recommends the use of RMAN for database backup and recovery in conjunction with Exadata. Essentially, using RMAN with Exadata is the same as using RMAN without Exadata.

To optimize the performance of incremental backups, the database can offload block filtering to Exadata Storage Server. This optimization is possible only when taking backups by using RMAN. The offload processing is done transparently without user intervention. During the offload processing, Exadata Storage Server filters out the blocks that are not required for the incremental backup in progress. Therefore, only the blocks that are required for the backup are sent to the database. This can significantly improve backup performance, especially if the proportion of changed blocks is low.

Hybrid Columnar Compression can assist to further improve backup performance. When a database uses Hybrid Columnar Compression, the number of blocks included in the backup is reduced compared to the same data in a database that uses a less space-efficient form of compression, or no compression at all. The reduced data size associated with Hybrid Columnar Compression results in smaller and ultimately quicker backups.

General Recommendations for RMAN

Use RMAN to back up and recover databases on Exadata.

- Use RMAN incremental backups and block change tracking.
- Use an external RMAN recovery catalog repository.
- Set `DB_RECOVERY_FILE_DEST_SIZE` to bound the space used in the Fast Recovery Area.

ORACLE

The following recommendations apply when using RMAN in conjunction with Exadata:

- **Use RMAN incremental backups and block change tracking:** Block change tracking helps with fast incremental backups by not scanning blocks that have not changed. Also, when performing incremental backups of databases on Exadata, additional block inspection is offloaded from the database servers. Block change tracking provides the greatest benefit for databases, where fewer than 20 percent of the blocks are changed between incremental backups. You may still benefit by using block change tracking with change rates greater than 20 percent. However, testing is recommended to ensure that backup times are reduced.
- **Use an external RMAN recovery catalog repository:** The RMAN recovery catalog should be hosted on a server outside Exadata. In practice, it is common to have a server that hosts the RMAN catalog, along with other management repositories, such as the Oracle Enterprise Manager repository and the Oracle Secure Backup catalog.

- **Set `DB_RECOVERY_FILE_DEST_SIZE` to bound the space used 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 enabled). It is important that you set the value of this parameter to less than the total free space in the disk group, taking into account 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.

Exadata Disk-Based Backup Strategy

Recommendations for Exadata disk-based database backups:

- Use a Fast Recovery Area
- Perform an initial level 0 (full) backup
- Perform periodic incremental level 1 backups
- Update your level 0 backup by applying the second to last level 1 backup

ORACLE

For disk-based database backups, Oracle recommends incremental backups to a Fast Recovery Area (FRA). The recommended strategy is outlined in the slide. Using this approach, you can achieve a good compromise between the time and effort required during backup and recovery operations. The approach also efficiently manages the amount of storage required for backups.

The initial configuration procedure for Exadata creates a disk group for the FRA, and the initial size of the disk group is determined by the selections made in the Exadata configuration tool.

Disk-Based Backup Recommendations

- Fast Recovery Area (FRA) configuration:
  - Default: FRA disk group striped across all available Exadata Storage Servers, along with data disk groups
    - High availability with the best throughput
  - Alternative: FRA disk group and data disk groups on separate Exadata Storage Servers
    - Separation of data and backups
    - Possibility of reduced throughput
- Additional RMAN recommendations:
  - Instances and channels:
    - Initially, run RMAN across all database instances with two RMAN channels per instance.
    - Configure up to eight RMAN channels per instance, if required.
  - Configure an Oracle Service to use as the RMAN target.

ORACLE

By default, the installation and configuration procedure for Exadata yields a disk group for the Fast Recovery Area (FRA), which is spread across all the disks in all the Exadata Storage Servers. In this configuration, the FRA shares each disk with disk groups containing database data files. By default, all the disk groups are created with at least `NORMAL` ASM redundancy (two-way mirroring). The result is that Exadata can tolerate the loss of an entire Exadata Storage Server and still maintain full data integrity.

Where customers desire a separation between database files and the FRA, the recommended alternative is to configure the disk group containing the FRA on a dedicated set of Exadata Storage Servers and to configure the required data disk groups on separate storage servers. This configuration means that the I/O for a single workload can no longer benefit from being striped across all the storage servers. However, it also provides additional protection if multiple simultaneous failures affect either the database or the FRA. Multiple simultaneous failures to both the database and FRA can still result in data loss.

Another strategy is to purchase additional high capacity Exadata Storage Servers or an Exadata Storage Expansion Rack specifically to store the FRA. This allows your databases to leverage the full capacity of the Exadata Storage Servers.

In addition to the general recommendations for RMAN in conjunction with Exadata, Oracle recommends the following configuration for disk-based backups:

- Testing shows that optimal backup rates are achieved by using all database instances with between two and eight RMAN channels per instance. Using this configuration, less than five percent of CPU is typically used for backup operations. Start by running RMAN on one database instance using two RMAN channels. If greater backup throughput is required, use additional instances with two more channels per instance. Finally, if required, configure up to eight channels per instance.
- Configure an Oracle Service to run against specific database instances in the cluster. The service is used by RMAN to automatically spread the backup load evenly among the target instances offering the service.
- For incremental backups, use the RMAN `BACKUP` command option `FILESPERSET 1`. This backup option leads to faster single file restore operations.
- Set the initialization parameter `_file_size_increase_increment=2143289344` to optimize the space used when incremental (level 1) backups are written to the FRA.
- 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. In releases before 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. 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.

Disk-Based Backup on Non-Exadata Storage

- Oracle Zero Data Loss Recovery Appliance and ZFS Storage Appliance provide high-performance and cost-effective alternatives that:
  - Are tested and validated with Exadata
  - Can connect directly to the Exadata InfiniBand network
- Other NAS and NFS-based options are feasible.
  - Thorough testing is required to ensure acceptable performance.
- Connection to a SAN requires an intermediate server.
  - Exadata does not support direct SAN connection.
  - Performance is likely to be limited by the intermediate server.

ORACLE

Oracle Zero Data Loss Recovery Appliance and Oracle ZFS Storage Appliance provide excellent backup and recovery performance. Both solutions have been tested and validated in conjunction with Exadata, and both can connect directly to the Exadata InfiniBand network.

Oracle ZFS Storage Appliance is a Network Attached Storage (NAS) platform that can be used as a target for RMAN backups on Exadata. For more information, see the Oracle ZFS Storage Appliance for Backup home page at http://www.oracle.com/storage/nas/zs-backup/index.html, and *Backup and Recovery Performance and Best Practices using Sun ZFS Storage Appliance with Oracle Exadata Database Machine* at http://www.oracle.com/technetwork/database/features/availability/maa-wp-dbm-zfs-backup-1593252.pdf.

Oracle Zero Data Loss Recovery Appliance is the first engineered system designed specifically for database protection. It delivers continuous protection for critical databases while offloading backup processing from production servers to minimize overhead. For more information, see the Oracle Zero Data Loss Recovery Appliance home page at http://www.oracle.com/engineered-systems/zero-data-loss-recovery-appliance/index.html.

Other options using Network Attached Storage (NAS) and Network File System (NFS) technologies are also feasible. However, users should thoroughly test the proposed solutions to ensure that their recovery time objectives can be met.

Users wishing to host backups using Storage Area Network (SAN) technologies must factor in the requirement for an intermediate server to act as an iSCSI or NFS gateway between Exadata and the SAN storage. An intermediate server is required because adding Fibre Channel adapters to the Exadata servers is not supported. Users should also note that the intermediate server is likely to limit the backup and recovery throughput that can be achieved by using a SAN.

Tape-Based Backup Strategy

For tape-based database backups:

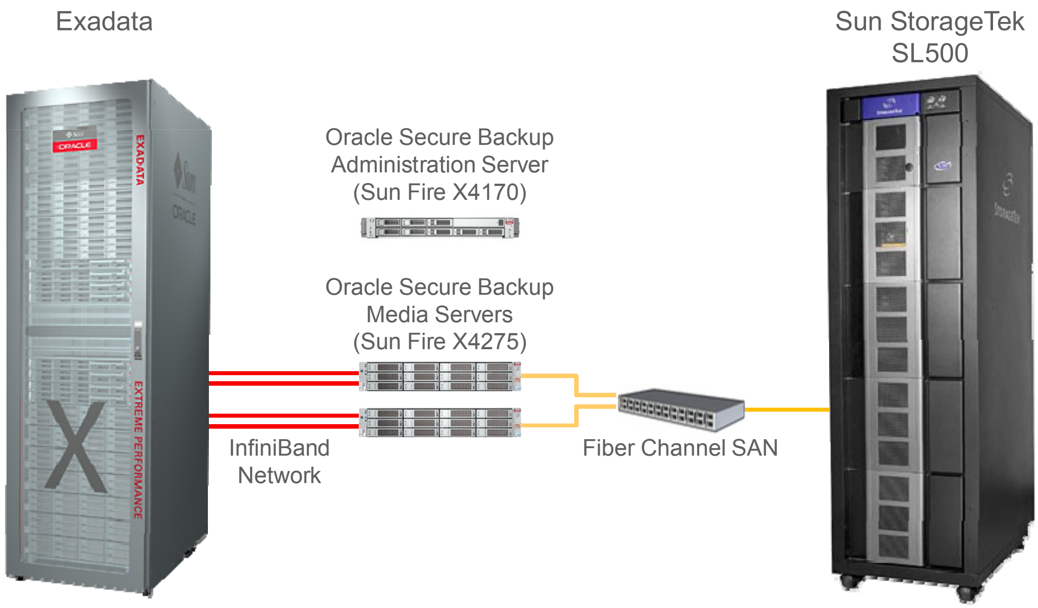
- Use media management software that is integrated with RMAN, such as Oracle Secure Backup.
- Perform periodic level 0 (full) database backups.
- Perform more frequent cumulative level 1 backups, and also back up the Oracle Secure Backup catalog.
- You can also backup you archived log files even more frequently for additional protection.

For tape-based database backups, Oracle recommends the use of Oracle Secure Backup or other media management software that is integrated with RMAN.

The recommended backup strategy is to perform periodic (weekly) level 0 (full) database backups. In addition, more frequent (daily) cumulative level 1 backups should be taken along with a backup of the Oracle Secure Backup catalog if the media management software is used.

In addition, many users find it beneficial to perform additional backups of archived redo log files several times in each cumulative backup cycle. For example, they may perform a cumulative level 1 backup on a daily basis with six-hourly backups of their archived redo log files.

Tape-Based Backup Architecture



The slide illustrates the recommended reference architecture for a tape-based backup configuration using Oracle Secure Backup. The key recommendations associated with the configuration follow:

- Connect the media servers to Exadata by using the high-performance InfiniBand network. The InfiniBand connection to each media server should be bonded for high availability. Alternatively, you can use an Ethernet network between the media servers and Exadata. However, if you use this configuration, be aware that the network is likely to be the constraining factor for backup and recovery performance.
- Consider the bandwidth capability of the InfiniBand connection from Exadata to the media server, compared to the SAN links between the media server and the tape library. Testing shows that data rates of approximately 3 GB per second can be sustained over an InfiniBand link. Therefore, you need to provide the appropriate number of SAN links and configure them to deliver the required throughput.
- Typically, backup performance is limited by tape drive throughput. Backup performance scales when you add more tape drives and RMAN channels. Allocate a sufficient number of tape drives so that the media servers can achieve their maximum backup and restore rates. Add tape drives until the bandwidth of the media servers is saturated.
- Start with at least two media servers. Add media servers if you have enough tape drives to keep them busy without saturating Exadata resources.

Tape-Based Backup Recommendations

- Media server to Exadata network configuration recommendation:
  - Use InfiniBand for the best backup rates:
    - Configure bonding for the media server InfiniBand interfaces.
    - Update OpenFabrics Enterprise Distribution on the media server.
    - Configure IP over InfiniBand connected mode for best performance.
    - Set the message transfer unit (MTU) size to 65520 for the InfiniBand interface.
    - Configure the media management software to use the InfiniBand network.
- Media server SAN configuration recommendation:
  - Configure persistent bindings for tape devices.

ORACLE

Media servers can be directly connected to the Exadata InfiniBand fabric by adding an InfiniBand Quad Data Rate (QDR) host channel adapter (HCA) to the media server. Use InfiniBand for the best backup rates, especially for large databases that require fast backup rates and low CPU overhead.

Follow these recommendations:

- For high availability, connect the HCA to two different Exadata InfiniBand leaf switches to eliminate the switch as a single point of failure. This provides transparent failover if connectivity is lost to one of the ports. Configure bonding of the media server InfiniBand (IB) interfaces. Use Active-Passive bonding.
- You must use an OpenFabrics Enterprise Distribution (OFED) version on the media server that is compatible with the version found in Exadata. For details, refer to My Oracle Support note 888828.1.
- Configure the InfiniBand interface to use IP over InfiniBand connected mode for best performance. On Linux, edit the `/etc/ofed/openib.conf` file so that it contains the entry `SET_IPOIB_CM=yes`. Restart the server to enable the setting. To verify the setting, check the contents of `/sys/class/net/ib0/mode` and `/sys/class/net/ib1/mode`. Both files should contain the entry `connected`.

- Set the message transfer unit (MTU) size to 65520 for the InfiniBand interface. On Linux, edit the `/etc/sysconfig/network-scripts/ifcfg-ib*` and `/etc/sysconfig/network-scripts/ifcfg-bond0` files so that they contain the entry `MTU=65520`. Verify the MTU setting by examining the output of the `ifconfig` command for the InfiniBand interfaces.
- To direct the backup and restore traffic over the InfiniBand network, configure the media management software to favor InfiniBand. Note that each media management software type has its own method of enabling this configuration. For example, 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 may require this configuration to be defined when the software is installed. See the media management software for information about how to direct traffic over a particular network.

For SAN-attached tape devices, configure persistent bindings so that 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 the media server software. Therefore, it is very important that your environment maintains consistent device addresses.

Persistent bindings are a part of the SAN infrastructure setup. Typically, persistent bindings are configured through the HBA or the operating system. The configuration steps will vary by platform and vendor. See My Oracle Support note 971386.1 for an example of creating persistent bindings for device attachments.

Connecting the Media Server by Using Ethernet

- Ethernet can be used if throughput is sufficient:
  - GigE: Expect up to 120 MB/sec from each interface.
  - 10gigE: Expect up to 1 GB/sec from each interface.
- Recommendations:
  - Use a dedicated backup network:
    - Configure dedicated network interfaces on each Exadata database server.
  - Use bonded network interfaces:
    - Configure LACP for maximum throughput:
      - Availability still maintained if one link is lost
      - Configuration required on media server, network switch, and database servers
    - Otherwise, use Active-Passive bonding for high availability:
      - Configure database servers and media servers
      - No specific switch configuration required

Alternatively, you can use a Gigabit Ethernet (GigE) or 10 Gigabit Ethernet (10gigE) network between Exadata and the media servers if throughput is sufficient.

Using a dedicated backup network eliminates any impact on the client access network. Often, a dedicated backup network is already in place. The maximum throughput with the GigE network is approximately 120 MB/sec for each interface. For 10gigE, the maximum throughput is approximately 1 GB/sec for each interface. Therefore, a Full Rack Exadata Database Machine can achieve throughput up to 960 MB/sec using GigE, or 8 GB/sec using 10 GigE, with a single dedicated network interface on each database server.

Higher throughput can be achieved by using Link Aggregation Control Protocol (LACP). LACP enables a bonded network interface to use both network channels simultaneously. To use LACP, you must configure a bonded network interface on each database server and set the `BONDING_OPTS` parameter to `mode=4` in the associated bonding configuration file (`ifcfg-bond1` for example). You must also configure LACP on the associated network switch and on the media server network interfaces. If LACP is not used, use Active-Passive bonded network interfaces to provide high availability in case of network interface failure.

Be aware that the ability to use a dedicated, bonded network interface for backup and recovery depends on the rest of the Exadata configuration. For example, there may not be sufficient network interfaces available if bonding is also used for the client access network. In this case, customers must balance the competing requirements.

Tape-Based Backup Recommendations

- Run RMAN across all the available database instances.
  – Create a Database Service that runs across the cluster:

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

  – Use the service name and SCAN address to connect RMAN:

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

- Allocate one RMAN channel per tape drive.
- Configure IORM and DBRM to control resource allocation between backups and application workloads.

ORACLE

Further recommendations:

- For your RAC databases, configure RMAN to run across all the database instances. It is recommended that you create a Database Service to run across the RAC cluster. Then when running RMAN, use the service name and SCAN address in the connect string for the RMAN target as shown in the slide.
- Allocate one RMAN channel per tape drive. A single RMAN channel can stream data at a rate of approximately 750 MB/sec from Exadata to the media server. 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.
- If resources must be prioritized between application workloads and backups, then configure I/O Resource Manager (IORM) and Database Resource Manager (DBRM). This is more likely to be required in cases where time-consuming backups to large databases must run at the same time as production application workloads.

Hybrid Backup Strategy

A hybrid backup strategy combines the disk-based and tape-based backup approaches.

- For example, disk-to-disk-to-tape
  - Initially, perform all backups to disk:
    - Readily available with high-performance access
  - Over time, copy backups to tape:
    - Long-term off-site storage
    - Relatively cheap backup media
- Follow recommendations for each approach in each phase.

A hybrid backup strategy combines the disk- and tape-based backup approaches.

For example, an approach known as disk-to-disk-to-tape (D2D2T) is a reasonable hybrid approach. Using this approach, all backups are initially performed to disk, providing high performance and easy availability. Then, over time the backups are copied to tape to provide log-term off-site storage using relatively cheap media.

If you choose to implement a hybrid backup approach, you should follow the recommendations for disk- or tape-based backups that are relevant in each different phase.

Restore and Recovery Recommendations

- Restore into existing data files, if possible.
  - Restore performance is better.
- Restore using all database instances.
- Recommended number of RMAN channels:
  - For disk-based restoration, use between two and eight RMAN channels per database instance.
  - For tape-based restoration, set the number of RMAN channels based on the total number of tape drives.

ORACLE

Higher restore rates are achieved when avoiding the overhead of initial data file allocation. If existing files are present before database restoration, do not delete the data files and perform database restoration into the existing files to take advantage of this optimization.

Create a restore service across all the database instances and use between two and eight RMAN channels per database instance for disk-based restoration. For tape-based restoration, the number of channels should be set to the number of tape drives.

Backup and Recovery of Database Machine Software

- Database Server software
  - Perform file system–level backup and recovery:
    - Use your chosen file system backup management software or Oracle Secure Backup.
    - Copies of the Oracle Cluster Registry are automatically maintained on the Database Server file system and should be included in Database Server file system backups.
- Exadata Storage Server software
  - File system level–backups are not recommended:
    - System areas are mirrored.
      - Use CellCLI commands to recover if one system disk fails.
    - Use the Exadata Software Rescue Procedure if both system disks fail simultaneously.
      - The rescue procedure uses a built-in USB flash drive.

ORACLE

You can use file system–level backup and recovery techniques for the database servers. Use your chosen file system backup management software and infrastructure or use Oracle Secure Backup, if desired. Note that a backup of the Oracle Cluster Registry, which also contains Voting Disk information, is automatically maintained on the file system of the first database server and must be included in your database server file system backups. The default location for the Oracle Cluster Registry backup is <Grid_Home>/cdata/<Cluster_Name>, where:

- <Grid_Home> is the location of the Grid Infrastructure software as specified during the initial configuration of Exadata. The default <Grid_Home> location is `/u01/app/11.2.0/grid`.
- <Cluster_Name> is the name of your cluster. This is the same as the Database Machine Prefix specified during the initial configuration of Exadata.

File system–level backups are not recommended for Exadata cell software. Exadata Storage Server maintains mirrored system areas on separate physical disks. In the rare event that both system disks fail simultaneously, you must use the rescue functionality provided on the CELLBOOT USB flash drive that is built into every Exadata Storage Server.

Quiz

How many RMAN channels should you use for tape-based backups?

a. Two per database instance
b. Four per database instance
c. One per tape drive

Answer: c

Summary

In this lesson, you should have learned how to:

- Explain how RMAN backups are optimized using Exadata Storage Server
- Describe the recommended approaches for disk- and tape-based backups of databases on Exadata
- List the recommended best practices for backup and recovery on Exadata

Additional Resources

- My Oracle Support notes
  - Database Machine and Exadata Storage Server 11*g* Release 2 (11.2) Supported Versions
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=888828.1
  - OSB - Create persistent bindings for device attachments on OEL
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=971386.1
- Documentation and white papers
  - Backup and Recovery Performance and Best Practices for Exadata Database Machine - Oracle Database 11.2.0.2 and later
    http://www.oracle.com/technetwork/database/features/availability/maa-tech-wp-sundbm-backup-11202-183503.pdf
  - Backup and Recovery Performance and Best Practices using Sun ZFS Storage Appliance with Oracle Exadata Database Machine
    http://www.oracle.com/technetwork/database/features/availability/maa-wp-dbm-zfs-backup-1593252.pdf

Practice 20 Overview: Using RMAN Optimizations for Exadata

In these practices, you will examine the backup and recovery optimizations that are enabled when Oracle Recovery Manager (RMAN) is used in conjunction with Exadata storage.

ORACLE

21

Exadata Database Machine Maintenance Tasks

ORACLE

Objectives

After completing this lesson, you should be able to perform the following Exadata maintenance tasks:

- Powering Exadata on and off
- Safely shutting down a single Exadata Storage Server
- Replacing a damaged physical disk on a cell
- Replacing a damaged flash card on a cell
- Moving all disks from one cell to another
- Using the Exadata cell software rescue procedure

Exadata Maintenance: Overview

- Maintaining Exadata is similar to maintaining any other clustered Oracle Database environment.
- Exadata–specific tasks outlined in this lesson:
  - Powering Exadata on and off
  - Safely shutting down a single Exadata Storage Server
  - Replacing a damaged physical disk on a cell
  - Replacing a damaged flash card on a cell
  - Moving all disks from one cell to another
  - Using the Exadata cell software rescue procedure
- Additional references:
  - *Exadata Database Machine Maintenance Guide*
  - My Oracle Support

ORACLE

In many respects, maintaining Exadata is similar to maintaining any other clustered Oracle Database environment. The procedures for maintaining Oracle clusterware, ASM, and Real Application Clusters (RAC) are essentially the same on Exadata as they are on other platforms. The main difference is the reference to Exadata cell objects.

This lesson focuses on a series of Exadata–specific maintenance tasks that administrators are most likely to encounter. Additional, less common maintenance tasks are also documented in the *Exadata Database Machine Maintenance Guide*. Administrators should also consult My Oracle Support for notes on other maintenance issues.

Note: Patching guidelines for Exadata are considered in a separate lesson later in the course.

Powering Exadata Off and On

- Power-off sequence:
  1. Database servers

     ```
     # <Grid_Home>/bin/crsctl stop cluster
     # shutdown -h now
     ```

     – Ensure that all database servers are shut down before proceeding.
  2. Exadata Storage Servers

     ```
     # shutdown -h now
     ```

     – Ensure that all storage servers are shut down before proceeding.
  3. Rack, including network switches
- Power-on sequence:
  1. Rack, including network switches
     – Apply power for a few minutes before proceeding.
  2. Exadata Storage Servers
     – Check that all cells are running before proceeding.
  3. Database servers

ORACLE

The slide lists the recommended sequence for powering Exadata off and on in non-emergency situations.

When performing either sequence, ensure that each step is completely finished before moving to the next step. Failure to execute steps in the proper order may result in Exadata not functioning correctly.

When powering on the Exadata Storage Servers and the database servers, power can be applied by pressing the power button at the front of each server, or remotely using the ILOM interface for each server. The ILOM can be accessed using the Web console, the command-line interface (CLI), IPMI, or SNMP. For example, you can start a server from the ILOM CLI by executing the `start /SYS` command.

To power the rack on or off, use the switches located on the Power Distribution Units (PDUs), which are located at the rear of the rack.

Safely Shutting Down a Single Exadata Storage Server

- Safe shutdown sequence:
  1. Stop the cell services:

     ```
     CellCLI> ALTER CELL SHUTDOWN SERVICES ALL
     ```

     – The command checks to ensure that it is safe to stop the cell.
  2. Power off the storage server.
- Startup sequence:
  1. Start the storage server.
     – Cell services start automatically.
  2. Verify that all the grid disks are active:

     ```
     CellCLI> LIST GRIDDISK ATTRIBUTES name, asmmodestatus
     ```

ORACLE

In some maintenance scenarios, a single Exadata Storage Server must be shut down in isolation. For example, a hardware component, such as a flash memory card or a disk controller, may be indicating an intermittent fault, so the storage server must be shut down to replace the component while the rest of the system continues to support business transactions. In these cases, the desired result is that the Exadata environment continues to support processing activities without a substantial impact on system users.

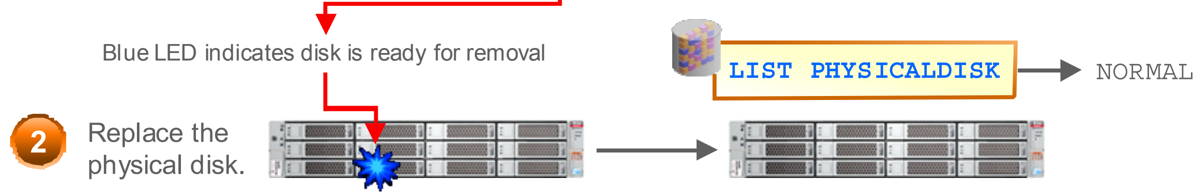
To safely and gracefully shut down a single Exadata Storage Server, use the commands shown in the slide. Note that the `ALTER CELL SHUTDOWN SERVICES ALL` command performs checks to ensure that it is safe to shut down the cell services. If it is not safe to do so, the command returns an error indicating the issue and the services are left running. This may occur if the cell is required to maintain disk group availability because other cells or disks are not available.

After the Exadata Storage Server is restarted, check that all the grid disks allocated to an ASM disk group display `asmmodestatus=ONLINE`. Grid disks that are not allocated to an ASM disk group should display `asmmodestatus=UNUSED`.

Replacing a Damaged Physical Disk

1 Determine the damaged disk.

```
CellCLI> LIST PHYSICALDISK DETAIL
         name:             20:5
         diskType:         HardDisk
         ...
         slotNumber:       5
         status:           failed
```



3 Monitor ASM to confirm the readdition of the disk.

```
SQL> SELECT NAME, STATE FROM V$ASM_DISK
SQL> SELECT * FROM GV$ASM_OPERATION
```

ORACLE

Replacing a physical disk, either a hard drive on a High Capacity cell or a flash drive on an Extreme Flash cell, due to a problem or failure, is probably the most likely hardware maintenance operation that an Exadata Storage Server may require. Because Exadata uses ASM redundancy, the procedure to replace a problem disk is quite simple.

The first step requires that you identify the problem disk. This could occur in several ways:

- For a failed physical disk or a physical disk with poor performance, the blue ready-to-remove LED is automatically switched on. For a physical disk with predictive failure status, the blue ready-to-remove LED is automatically switched on after all grid disks present on the physical disk are dropped from ASM and redundancy is fully restored. In both cases, the blue LED indicates that it is safe to remove and replace the affected disk.
- Hardware monitoring using ILOM may report a problem disk.
- If a disk fails, an Exadata alert is generated. The alert includes specific instructions for replacing the disk. If you have configured the system for alert notifications, the alert will be sent to the designated email address or SNMP target. The `LIST ALERTHISTORY` command could also be used to identify a failed disk.
- The `LIST PHYSICALDISK` command may identify a disk reporting an abnormal status. Even if the disk is still functioning, Exadata may detect a problem that may be a precursor to a disk failure and place the disk in a predictive failure state.

When a failed disk is detected, the Oracle ASM disks associated with the grid disks on the physical disk are automatically dropped with the `FORCE` option, and an Oracle ASM rebalance follows to restore the data redundancy. This process is known as proactive disk quarantine.

After you have identified the problem disk, you can replace it. When you remove the disk, you will get an alert. When you replace a physical disk, the disk must be acknowledged by the RAID controller before it can be used. This does not take a long time, and you can use the `LIST PHYSICALDISK` command to monitor the status until it returns to `NORMAL`.

The grid disks and cell disks that existed on the previous disk in the slot will be automatically recreated on the new disk. If these grid disks were part of an Oracle ASM disk group with `NORMAL` or `HIGH` redundancy, they will be added back to the disk group and the data will be rebalanced based on disk group redundancy and the `asm_power_limit` parameter.

Recreating the ASM disk and rebalancing the data may take some time to complete. You can monitor the progress of these operations within ASM. You can monitor the status of the disk as reported by `V$ASM_DISK.STATE` until it returns to `NORMAL`. You can also monitor the rebalance progress by using `GV$ASM_OPERATION`.

Safe Disk Removal

Sometimes a disk must be removed even though the system detects no fault with it:

```
CellCLI> ALTER PHYSICALDISK { disk_1 [ , disk_n ] } DROP FOR REPLACEMENT
```

- Disables the disk and switches on the blue ready-to-remove LED only if it is safe to remove the disk
- Requires Exadata release 11.2.3.3.0 or later

ORACLE

Sometimes a disk must be replaced proactively, even though its blue ready-to-remove LED is switched off. For example, the `CALIBRATE` command may identify a disk delivering abnormally low throughput or I/Os per second (IOPS).

However, removing the disk could lead to a disk group force dismount if the removed disk has the only copy of some pieces of data, or a cell crash due to loss of its OS if the disk is the only system disk.

To mitigate this, Exadata release 11.2.3.3.0 adds a new command. `ALTER PHYSICALDISK ... DROP FOR REPLACEMENT` checks whether the specified physical disks can be removed safely, and if so, the disks are disabled and the blue ready-to-remove LEDs are switched on. This command replaces checks that were previously required to ensure that disks could be removed safely. By using this command, disk replacements are simplified and problems are avoided.

In addition, `ALTER PHYSICALDISK ... REENABLE FORCE` reenables normal physical disks. This command can be used to reenable physical disks in situations where `ALTER PHYSICALDISK ... DROP FOR REPLACEMENT` was executed but the administrator does not want to remove the disks.

Replacing a Damaged Flash Card

1. Determine the damaged flash card.

```
CellCLI> LIST PHYSICALDISK DETAIL
         name:            FLASH_5_3
         diskType:        FlashDisk
         ...
         slotNumber:      "PCI Slot: 5; FDOM: 3"
         status:          failed
```

2. Power off the cell.
3. Replace the flash card.

4. Power on the cell.

ORACLE

Identifying a damaged flash module is similar to identifying a damaged physical disk. Hardware monitoring using ILOM or a drop in performance indicated by the `CALIBRATE` command may indicate a problem. If a failed FDOM is detected, an alert is generated.

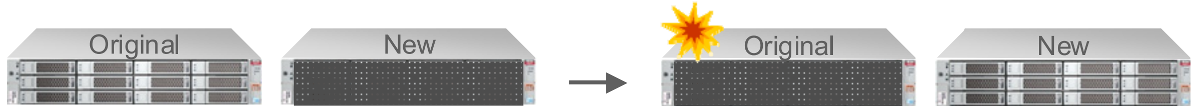
A bad flash module results in a decreased amount of flash memory on the cell. The performance of the cell is affected proportional to the size of flash memory lost, but the database and applications are not at risk of failure.

As shown in the slide, a damaged flash module can also be reported by using the `LIST PHYSICALDISK DETAIL` command. The `slotNumber` attribute shows the PCI slot and the FDOM number. In this example, the `status` attribute indicates a device failure.

Although the PCI slots in an Exadata Storage Server support hot replacement, it is recommended to power off the cell while replacing a damaged flash card.

After replacing the card and powering on the cell, no additional steps are required to recreate the Smart Flash Cache and Smart Flash Log areas on the new flash modules.

Moving All Disks from One Cell to Another



1. Back up the following files and directories:
   - Files: `/etc/hosts`, `/etc/modprobe.conf`
   - Directories: `/etc/sysconfig/network`, `/etc/sysconfig/network-scripts`
2. Deactivate the grid disks: `CellCLI> ALTER GRIDDISK ALL INACTIVE`
3. Shut down the cell.
4. Move the disks, flash cards, disk controller, and CELLBOOT USB flash drive from the original cell to the new cell.
   - Ensure that the system disks occupy the same first two slots.
   - Ensure that the flash cards occupy the same PCI slots.
5. Power on the new cell.
6. Check the operating system configuration.
7. Restart the cell services: `CellCLI> ALTER CELL RESTART SERVICES ALL`
8. Activate the grid disks: `CellCLI> ALTER GRIDDISK ALL ACTIVE`
9. Validate the configuration: `CellCLI> ALTER CELL VALIDATE CONFIGURATION`
10. Activate the ILOM for ASR.

ORACLE

You may need to move all drives from one storage server to another. This may be necessary when there is a chassis-level component failure, or when troubleshooting a hardware problem. To move the drives, perform the following steps:

1. If possible, back up `/etc/hosts`, `/etc/modprobe.conf`, and the files in `/etc/sysconfig/network` and `/etc/sysconfig/network-scripts`. This is mainly a precautionary step and is also useful if you want to move the disks back to the original chassis.
2. If possible, deactivate the grid disks by using `ALTER GRIDDISK ALL INACTIVE`. Also, ensure that the Oracle ASM `disk_repair_time` attribute is set sufficiently long enough so ASM does not drop the disks before the grid disks can be activated in another Exadata Storage Server.
3. Shut down the original server.
4. Move the hard disks, flash cards, disk controller card, and CELLBOOT USB flash drive to the new server.

   Caution: Ensure that the first two disks, which are the system disks, are in the same first two slots. Also ensure that the flash cards are installed in the same PCI slots. Failure to do so causes the Exadata Storage Server to function improperly.

5. Start the new server. The cell operating system should be automatically reconfigured to suit the new server hardware.
6. Check `/etc/hosts`, `/etc/modprobe.conf`, and the files in `/etc/sysconfig/network` and `/etc/sysconfig/network-scripts`. If any files are corrupted, restore them from your backup or manually re-create them using a good cell as a guide. Also, use the `ifconfig` command to confirm the hardware addresses for the Ethernet interfaces (`eth0`, `eth1`, and so on). Compare the Ethernet interface hardware addresses with the `HWADDR` value stored in the configuration files (`ifcfg-eth0`, `ifcfg-eth1`, and so on) in the `/etc/sysconfig/network-scripts` directory. If necessary, edit the configuration files so that the `HWADDR` value matches the corresponding `ifconfig` command output.
7. Restart the cell services by using `ALTER CELL RESTART SERVICES ALL`.
8. Activate the grid disks by using `ALTER GRIDDISK ALL ACTIVE`.
9. Validate the cell configuration by using `ALTER CELL VALIDATE CONFIGURATION`.
10. If required, activate the ILOM interface of the new cell for Auto Service Request (ASR).

If you are using ASM redundancy and the procedure is completed before the amount of time specified in the `DISK_REPAIR_TIME` ASM initialization parameter, then the ASM disks will be automatically brought back online and updated with any changes made during the cell outage.

Using the Exadata Cell Software Rescue Procedure

- Every Exadata Storage Server is equipped with a CELLBOOT USB flash drive to facilitate cell rescue:
  - Required if both system disks fail simultaneously or are corrupt
  - Should be used with extreme caution
- To perform cell rescue:
  1. Connect to the Exadata Storage Server using the console.
  2. Boot the cell, and as soon as you see the Oracle Exadata splash screen, press any key on the keyboard.
  3. From the boot options list, select the last option, `CELL_USB_BOOT_CELLBOOT_usb_in_rescue_mode`. Then press Enter.
  4. Select the rescue option, and proceed with the rescue.
  5. At the end of the rescue process, ensure that the cell boots from the system disks.
  6. Reconfigure the cell.

Exadata Storage Server maintains mirrored system areas on separate physical disks. If one system area becomes corrupt or unavailable, the mirrored copy is used to recover.

In the rare event that both system disks fail simultaneously, you must use the rescue functionality provided on the CELLBOOT USB flash drive that is built into every Exadata Storage Server. The rescue procedure may also be required to recover from a file system corruption or a corrupt boot area.

It is important to note the following when using the rescue procedure:

- Use extreme caution when using this procedure, and pay attention to the prompts. The rescue procedure can rewrite some or all of the disks in the cell. If this happens, you can lose the contents of those disks. You should use the rescue procedure only with assistance from Oracle Support Services.
- The rescue procedure does not destroy the contents of the data disks or the contents of the data partitions on the system disks unless you explicitly choose to do so during the rescue procedure.
- The rescue procedure restores the Exadata Storage Server software to the same release. This includes any patches that existed on the cell as of the last successful boot.

- The following is not restored using the rescue procedure:
  - Some cell configuration details, such as alert configurations, SMTP information, and administrator email address. Note that the cell network configuration is restored, along with SSH identities for the cell, and the `root`, `celladmin`, and `cellmonitor` users.
  - ILOM configurations. Typically, ILOM configurations remain undamaged even in case of Exadata software failures.
- The rescue procedure does not examine or reconstruct data disks or data partitions on the system disks. If you have data corruption on the grid disks, do not use the rescue procedure. Instead, use the database backup and recovery procedures.

The following rescue options are available for the rescue procedure:

- **Partial reconstruction recovery:** During partial reconstruction recovery, the rescue process re-creates partitions on the system disks and checks the disks for the existence of a file system. If a file system is discovered, the process attempts to boot. If the cell boots successfully, you use the CellCLI commands, such as `LIST CELL DETAIL`, to verify that the cell is usable. You must also recover any data disks, as appropriate. If the boot fails, you must use the full, original build recovery option.
- **Full, original build recovery:** This option rewrites the system area of the system disks to restore the Exadata software. It also allows you to erase any data on the data disks, and data partitions on the system disks.
- **Recreation of the CELLBOOT USB flash drive:** This option is used to make a copy of the CELLBOOT USB flash drive.

To perform a rescue by using the CELLBOOT USB flash drive:

1. Connect to the Exadata Storage Server by using the console.
2. Boot the cell, and as soon as you see the Oracle Exadata splash screen, press any key on the keyboard. The splash screen remains visible for only five seconds.
3. In the list of boot options, scroll down to the last option, `CELL_USB_BOOT_CELLBOOT_usb_in_rescue_mode`, and press Enter.
4. Select the rescue option, and proceed with the rescue.
5. When prompted at the end of the rescue process, proceed as follows:
   a. Choose to enter a shell. Do not select the reboot option at this point.
   b. Log in to the shell by using the rescue `root` password.
   c. Run the `reboot` command from the shell.
   d. During reboot, but before you see the Oracle Exadata splash screen, press F8 to access the boot device selection menu.
   e. Select the RAID controller as the boot device.
6. After a successful rescue, you must reconfigure the cell to return it to the pre-failure configuration. This may include items such as cell configuration parameters, IORM configuration and metric thresholds. If you selected to preserve the data when prompted by the rescue procedure, the cell disks are automatically imported during the rescue procedure. If you selected not to preserve the data, you should create new cell disks and grid disks.

Quiz

When shutting down an Exadata Database Machine, the Exadata Storage Servers must be shut down first.

a. True

b. False

Answer: b

Quiz

An Exadata Storage Server must be shut down to replace failed hardware components other than hard disk drives.

a. True

b. False

Answer: a

Although the flash memory cards inside Exadata Storage Server are hot-swappable, Oracle recommends that cells be shut down to replace hardware components inside the chassis.

Quiz

Which of the following statements are true about an Exadata Storage Server disk failure?

a. The associated ASM grid disks are automatically dropped and an ASM rebalance occurs to quickly restore redundancy.
b. The disk may be replaced without shutting down the storage server.
c. The storage server must be shut down to replace the disk.
d. Multiple ASM instances can participate in the rebalance operation of a single disk group.

ORACLE

Answer: a, b

Summary

In this lesson, you should have learned how to perform the following Exadata maintenance tasks:

- Powering Exadata on and off
- Safely shutting down a single Exadata Storage Server
- Replacing a damaged physical disk on a cell
- Replacing a damaged flash card on a cell
- Moving all disks from one cell to another
- Using the Exadata cell software rescue procedure

22

Patching Exadata Database Machine

Objectives

After completing this lesson, you should be able to:

- Describe how software is maintained on different Exadata components
- Locate the recommended patches for Exadata
- Describe the recommended patching process for Exadata
- Describe the characteristics of an effective test system

ORACLE

Patching and Updating: Overview

- There are three categories of software that must be maintained in Exadata:
  - Software and firmware on the Exadata Storage Servers
  - Software and firmware on the database servers
  - Software and firmware for other components
- Compatibility between different software must be maintained.
- Wherever possible, patches and updates are rolling in nature.
- Platinum Services for Exadata takes care of patching and updates:
  - Includes planning and deployment of four patches every year

ORACLE

There are three broad categories of software that must be maintained in Exadata: software and firmware in the Exadata Storage Servers, database servers, and other components (such as the InfiniBand switches). Compatibility between the software is vital for Exadata to function.

Wherever possible, Exadata patches and updates are rolling in nature, which means that they are applied in a manner that facilitates ongoing system availability by rolling the patch across the environment one server at a time, rather than affecting all the servers at once.

For customers who enable Oracle Platinum Services for Exadata, Oracle proactively plans and deploys patches four times every year.

The remainder of this lesson outlines additional recommendations regarding Exadata patching and updating.

Patching and Updating: Key Information Sources

- Exadata Database Machine and Exadata Storage Server Supported Versions: My Oracle Support note 888828.1:
  - News on latest patches and software releases
  - Version compatibility and interoperability details
  - Historical information on patches and software releases
- Exadata Critical Issues: My Oracle Support note 1270094.1:
  - Up to date list of critical issues, fixes, and workarounds
- Use Exachk for latest recommendations:

Exachk Report:

FAIL	Patch Check	System is exposed to Exadata Critical Issue DB31	All Homes	View

Exadata Critical Issues (MOS note 1270094.1)

DB31	ASM 12.1.0.2	Bug 21281532 - ASM rebalance interrupted with errors ORA-600 [kfdAtbUpdate_11_02] and ORA-600 [kfdAtUnlock00].	Fixed in 12.1.0.2.11. See Document 2031709.1 for additional details.

ORACLE

Important information relating to Exadata patches and updates is maintained in My Oracle Support note 888828.1. This note is constantly updated with new information. Likewise, a log of Exadata Critical Issues, including fixes and workarounds, is maintained in My Oracle Support note 1270094.1. Oracle recommends that all Exadata customers should sign up for an automated alert when either of these notes change.

You should also use the Exachk utility to list what patches and updates are available for your system. Exachk also provides a report that outlines your exposure to known critical issues. An example is shown in the slide. To ensure you have access to the latest recommendations, use My Oracle Support note 1070954.1 to update to the latest version of Exachk.

Maintaining Exadata Storage Server Software

- Exadata Storage Server patches:
  - Are complete software images
  - Maintain consistency across all cell components
  - Are released independently from Oracle Database patches
  - Have dependencies on other software and firmware versions
- Most Exadata Storage Server patches can be applied while databases remain functional.
- Firmware levels in Exadata Storage Server are maintained automatically.
- No additional software, RPMs or otherwise, must be installed on Exadata Storage Server.

ORACLE

Patches for Exadata Storage Server are provided as complete software images that contain updates to the Linux operating system, cell server software, InfiniBand software, and component firmware. This ensures that all the software and firmware components on an Exadata Storage Server remain consistent with each other.

Exadata Storage Server patches are supplied independent of Oracle Database Server patches. However, an Exadata Storage Server patch may require a specific Oracle Database Server patch level, or database server firmware or operating system version. Details are provided in the patch documentation and in My Oracle Support note 888828.1.

Exadata Storage Server patches can generally be applied in a rolling manner, while the databases continue to run. Patches may also include instructions for parallel (non-rolling) installation on multiple Exadata Storage Servers.

Exadata Storage Server automatically maintains firmware within the server. Firmware levels are periodically checked while Exadata Storage Server is running, and the correct firmware is automatically applied to components when the server restarts. Disk firmware is also automatically updated when a disk is replaced.

Oracle recommends that no additional software, whether RPMs or otherwise, should be installed on Exadata Storage Server.

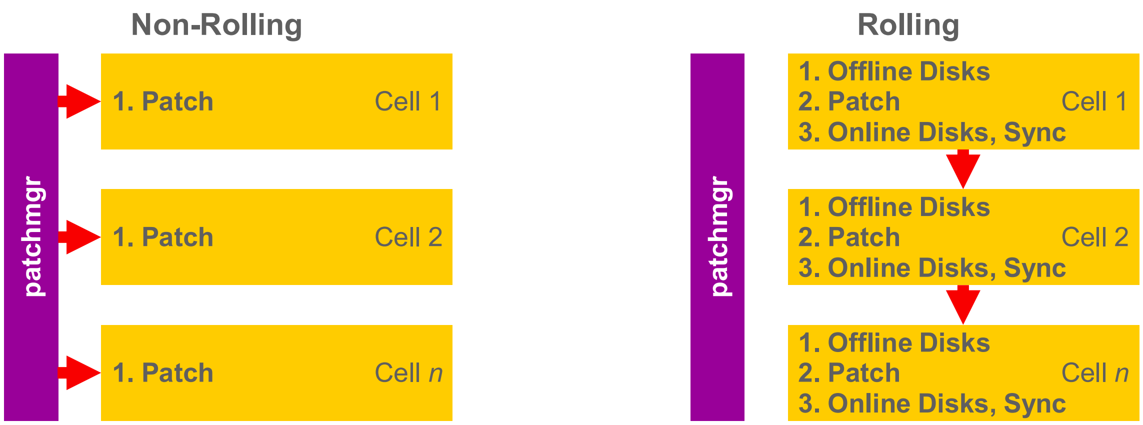
Using `patchmgr` to Orchestrate Storage Server Patching

- Download and stage the patch on one of the database servers
- Check prerequisites using:

```
# cd <patch_dir>
# ./patchmgr -cells <cell_group> -patch_check_prereq
```

- Apply patch using:

```
# ./patchmgr -cells <cell_group> -patch [-rolling]
```



ORACLE

Exadata Storage Server patches are typically installed using a script supplied with the patch called `patchmgr`. They may be installed in one of two ways: rolling and non-rolling:

- **Rolling:** A rolling patch is performed sequentially on each storage server. The rolling patch process relies on ASM redundancy to maintain disk group availability while each storage server is updated and restarted. Databases remain operational while all of the storage servers are patched. Each storage server can take up to approximately two hours to patch. Therefore, patch application on a Full Rack Database Machine can take more than a day to complete.
- **Non-rolling:** A non-rolling patch is applied to all storage servers simultaneously with databases offline. This approach requires down time, but may be preferred if a scheduled maintenance window exists for the system. Because all of the storage servers are patched in parallel, the whole patch make take approximately two hours to complete.

In either mode, a single invocation of `patchmgr` can be used to patch all of the Exadata Storage Servers.

In all cases, ensure that you follow the patch-specific instructions for using `patchmgr`.

Maintaining Database Server Software

- You can patch and update the database server software as you would for an Oracle Database server outside of Exadata.
  - Oracle Database patches applied by OPatch:
    - Double-check compatibility of patches with Exadata.
    - Oracle supplies Bundle Patches for Exadata:
      - Periodic bundling of database patches recommended for Exadata
  - Operating system and firmware updated using regular channels:
    - Use the Exadata-specific channel on ULN
    - Check to maintain consistency with InfiniBand (OFED) software
- Check Exadata Storage Server patches for database server firmware and operating system updates.

In essence, you can patch and update the database server software by using the same techniques and approaches as you would for an Oracle Database server outside of Exadata.

OPatch is used to apply and manage Oracle Database patches. You can raise a service request with Oracle Support to verify the compatibility of patches with Exadata. In addition to regular database patches and updates, Oracle supplies periodic Bundle Patches for Exadata. These bundle a series of database patches that are recommended for use in conjunction with Exadata. Oracle recommends that you apply Bundle Patches as part of your patching routine.

Operating system and firmware patches must be managed by using standard patching practices. Starting with Exadata version 11.2.3.1.0, you can use the Exadata-specific channel on the Unbreakable Linux Network (ULN) to get updates for database servers running Oracle Linux. My Oracle Support note 888828.1 lists current constraints and requirements for operating system and firmware patches. The main consideration is to ensure that any patches or updates maintain consistency with the OFED software version, which underpins the InfiniBand network.

Some Exadata Storage Server patches also include database server firmware and operating system updates. This may occur, for example, when the firmware for the InfiniBand HCA is updated on Exadata Storage Server and the update must also be applied to the database servers.

Assisted Patching Using OPlan

- OPlan is a utility that provides step-by-step patching instructions that are specific to the target environment:
  - Can create instructions for Apply and Rollback
- OPlan works in conjunction with Exadata recommended Bundle Patches:
  - Starting with 11.2.0.2 Bundle Patch 2
- Using OPlan:
  - As the Oracle software owner:
    - Download the Bundle Patch to a local directory
    - Set the `$ORACLE_HOME` environment variable and execute:

```
$ORACLE_HOME/oplan/oplan generateApplySteps <bundle patch location>
```

    - Locate the customized patch installation instructions at:

```
$ORACLE_HOME/cfgtoollogs/oplan/<TimeStamp>/InstallInstructions.txt
```

ORACLE

OPlan is a utility that assists the patch installation process by providing step-by-step patching instructions specific to your environment. OPlan works in conjunction with Exadata recommended Bundle Patches.

Without OPlan, administrators must read the patch documentation and translate the generic patch installation instructions into specific commands suitable for the target environment. For example, references to the Oracle home directory or the path to the patch directory must be translated into the appropriate directory paths. Also, administrators may be required to choose between several options and evaluate the configuration of the environment to generate the right set of commands.

OPlan assists administrators by automatically analyzing the environment and collecting the required configuration information. Using this information, OPlan generates a set of instructions and commands, which are customized specifically for the target configuration. It can create the instruction required to apply a patch and also the instruction required to perform a patch rollback.

The OPlan utility can be downloaded from My Oracle Support using patch number 11846294. OPlan supports all Exadata recommended Bundle Patches from release 11.2.0.2 Bundle Patch 2. See My Oracle Support note 1306814.1 for further information about OPlan.

Assisted Patching Using the DB Node Update Utility

- The DB Node Update Utility automates all the steps and checks to upgrade Exadata database servers to a new Exadata software release:
  - Updates the database server operating system
    - Does not update Oracle Database or Grid Infrastructure
  - Can perform upgrades and rollbacks
  - Compliments the Exadata Storage Server patching procedure and Oracle Database patching procedures
- The DB Node Update Utility supports the following:
  - Database servers with Oracle Linux 5.5 and later
  - Upgrades from release 11.2.2.4.2 and later
  - Rollbacks to release 11.2.2.4.2 and later
  - “Known issues” for upgrades to release 11.2.3.2.1 and later
- See My Oracle Support note 1553103.1 for details.

The DB Node Update Utility (`dbnodeupdate.sh`) automates all the steps and checks to upgrade Exadata database servers to a new Exadata software release. Specifically, the utility updates the database server operating system to match the requirements for a corresponding Exadata software release. It can also perform rollback operations. It includes the latest best practices and performs backups before patching where possible. The utility also includes known issue workarounds for upgrades starting with Exadata software release 11.2.3.2.1.

The utility does not upgrade Oracle Database software. However, it provides database patching recommendations where required and it relinks the Oracle binaries after updating the operating system.

The DB Node Update Utility compliments the Exadata Storage Server patching procedure and the Oracle Database software patching procedures.

By default, the utility works interactively on a single database server. However, multiple database servers can be updated at the same time by running the utility in parallel across them by using the `dcli` utility. However, it is recommended to first perform an upgrade on one server and then do the remaining servers in parallel.

The utility has different requirements and modes of operation depending on the source and target Exadata software releases in question. See My Oracle Support note 1553103.1 for details and examples of various different upgrade scenarios.

Using `patchmgr` to Orchestrate Database Server Patching

Starting with updates to Exadata release 12.1.2.2.0, Exadata database servers can be updated using `patchmgr`:

- Target servers must run Exadata release 11.2.2.4.2, or later
- `patchmgr` uses `dbnodeupdate.sh` on each node
- Multiple servers can be patched in one operation
  - But `patchmgr` cannot patch the server that it is running on
- Usage:
  - Download and stage `patchmgr` on one of the database servers
    - Check My Oracle Support note 1553103.1 for details
  - Check prerequisites using:

```
# ./patchmgr -dbnode <dbnode_list> -dbnode_precheck
  -dbnode_loc <patch_loc> -dbnode_version <version>
```

  - Apply patch using:

```
# ./patchmgr -dbnode <dbnode_list> -dbnode_upgrade
  -dbnode_loc <patch_loc> -dbnode_version <version>
```

ORACLE

Starting with updates to Exadata release 12.1.2.2.0, Oracle Exadata database servers, including virtualized servers, can be updated, rolled back, and backed up using `patchmgr`. To use this method, the database servers must be running Exadata software release 11.2.2.4.2, or later.

For Exadata database server patch orchestration, `patchmgr` runs the DB Node Update Utility (`dbnodeupdate.sh`) on each database server. You can still run `dbnodeupdate.sh` in standalone mode, but by performing the updates by running `patchmgr`, you can run a single command to update multiple nodes at the same time. Similar to storage server patching, `patchmgr` can update the nodes in a rolling or non-rolling fashion.

You must run `patchmgr` from a database server that will not be updated itself. That server must be updated later by running `patchmgr` from a server that has already been updated, or by running `dbnodeupdate.sh` on the server in question.

Starting with release 12.1.2.2.0, a new `dbserver.patch.zip` file is available for running `dbnodeupdate.sh` from `patchmgr`. Unzip the `dbserver.patch.zip` file and run `patchmgr` from there. Check My Oracle Support note 1553103.1 for the latest release of `dbserver.patch.zip`.

In all cases, ensure that you follow any patch-specific instructions for using `patchmgr`.

Maintaining Other Software

- Other components in Exadata that have software or firmware:
  - InfiniBand switches
    - Typically applied using `patchmgr`
  - Power Distribution Units (PDUs)
  - Keyboard, Video, and Mouse (KVM) switch
  - Cisco Ethernet switch
- Always refer to My Oracle Support note 888828.1 for Exadata-specific requirements.

The software and firmware on all components must always be maintained in accordance with the guidelines published in My Oracle Support note 888828.1. In particular, this is critical for the InfiniBand switches. For other components, there may be less urgency to keep up with the latest patches.

Recommended Patching Process

1. Review the patch documentation (README file).
   – Read and understand it before proceeding.
2. Validate the patch installation on a test system.
   – Run Exachk before and after patch application.
   – Automate the patch application steps, where possible.
   – Test the fallback procedure.
3. Validate the patch functionality on a test system.
   – Verify that the patch provides the desired functionality.
   – Evaluate system performance.
4. Apply the patch in production.
   – Run Exachk before and after patch application.
   – Make sure all cells are healthy before applying the patch.
   – Evaluate system performance.

ORACLE

No two patches are the same. This adage is true whether or not you are using Exadata. Therefore, you should always follow a methodical patching process. Following is an outline of the recommended patching process for Exadata:

1. Review the patch README file for known issues, patch installation and de-installation instructions, and special notes.
2. Validate the patch installation on a proper test system:
   - Run Exachk to verify the environment before and after patch application.
   - Where possible, automate the steps to reduce human error.
   - Test the fallback procedures in case the patch must be rolled back.
3. Validate the patch functionality on the test system.
   - If applying the patch to address a specific issue, verify that the patch provides the desired functionality.
   - Verify that there are no performance, availability, or operational regressions. Test using a workload that is representative of the production workload. Compare metrics observed during the test with baseline metrics observed in production. Real Application Testing may be used to replay a production workload on the test system. Automatic Workload Repository (AWR) and SQL Performance Analyzer may be used to assess performance on the test system.

4. Apply the patch on the production system.
   - Run Exachk to verify the environment before and after patch application.
   - Make sure that all the Exadata Storage Servers are functioning correctly. This is especially important because an unnoticed storage server fault that does not seem to impact Exadata operations may result in disk groups being taken offline and databases terminating abruptly when another storage server is taken offline for patching.
   - Compare the metrics captured before and after patch application to evaluate system performance. SQL Performance Analyzer can be used to assess performance improvement or regression resulting from the patch.

Refer to the lesson titled *Other Useful Monitoring Tools* and My Oracle Support note 1070954.1 for details regarding Exachk.

Test System Recommendations

An effective test system:

- Is an exact replica of the production Exadata system
- Is not used for any other purpose during testing
- Contains a full copy of the production data set with identical statistics
- Is able to mimic production transaction volumes and concurrency
- Is able to compare workload and performance metrics on the test system with metrics collected from the production system
- Includes the operational and availability test suites used in production

ORACLE

The following list describes the recommendations for an effective Exadata test environment:

- A test system should be an exact replica of the primary Exadata system, including any standby systems for high availability and disaster recovery.
- When the test system is shared for other purposes, such as development, exclusive time should be allocated to perform patch validation.
- The test system should contain a full copy of the production data set with identical statistics.
- A workload framework, such as Real Application Testing, should be used to mimic production transaction volumes and concurrency.
- AWR data should be collected from the production database to compare against AWR data collected on the test system.
- Any operational and availability test suites used in production should also be ported to the test environment.

Quiz

Which of the following statements is a broad overview of the recommended approach for patching and updating Exadata?

a. All updates for Exadata are specific to Exadata, and these are the only updates that should be applied to Exadata.

b. Use Unbreakable Linux Network (ULN) to update Linux on database servers and Exadata Storage Servers, and apply specific database and Exadata cell updates for the remaining software.

c. Use Exadata-specific update bundles for the Exadata Storage Servers, and use normal database updating practices for the database servers.

ORACLE

Answer: c

Summary

In this lesson, you should have learned how to:

- Describe how software is maintained on different Exadata components
- Locate the recommended patches for Exadata
- Describe the recommended patching process for Exadata
- Describe the characteristics of an effective test system

ORACLE

Additional Resources

- Lesson demonstrations
  - Exadata Storage Server Rolling Patch Application
    http://apex.oracle.com/pls/apex/f?p=44785:24:5854243986926427::NO:24:P24_CONTENT_ID,P24_PREV_PAGE:5523,29
  - Exadata Storage Server Patch Rollback
    http://apex.oracle.com/pls/apex/f?p=44785:24:5854243986926427::NO:24:P24_CONTENT_ID,P24_PREV_PAGE:5522,29
- My Oracle Support notes
  - Exadata Database Machine and Exadata Storage Server Supported Versions
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=888828.1
  - Exadata Critical Issues
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1270094.1

Additional Resources

- My Oracle Support notes
  - Oracle Software Patching with OPLAN
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1306814.1
  - Exadata Database Server Patching using the DB Node Update Utility and patchmgr
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1553103.1
  - Oracle Exadata Database Machine exachk or HealthCheck
    https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1070954.1

ORACLE

23

Exadata Database Machine Automated Support Ecosystem

ORACLE

Objectives

After completing this lesson, you should be able to:

- Describe the Auto Service Request (ASR) function and how it relates to Exadata
- Explain the implementation requirements for ASR
- Explain the ASR configuration process
- Describe Oracle Configuration Manager (OCM) and how it relates to Exadata

ORACLE

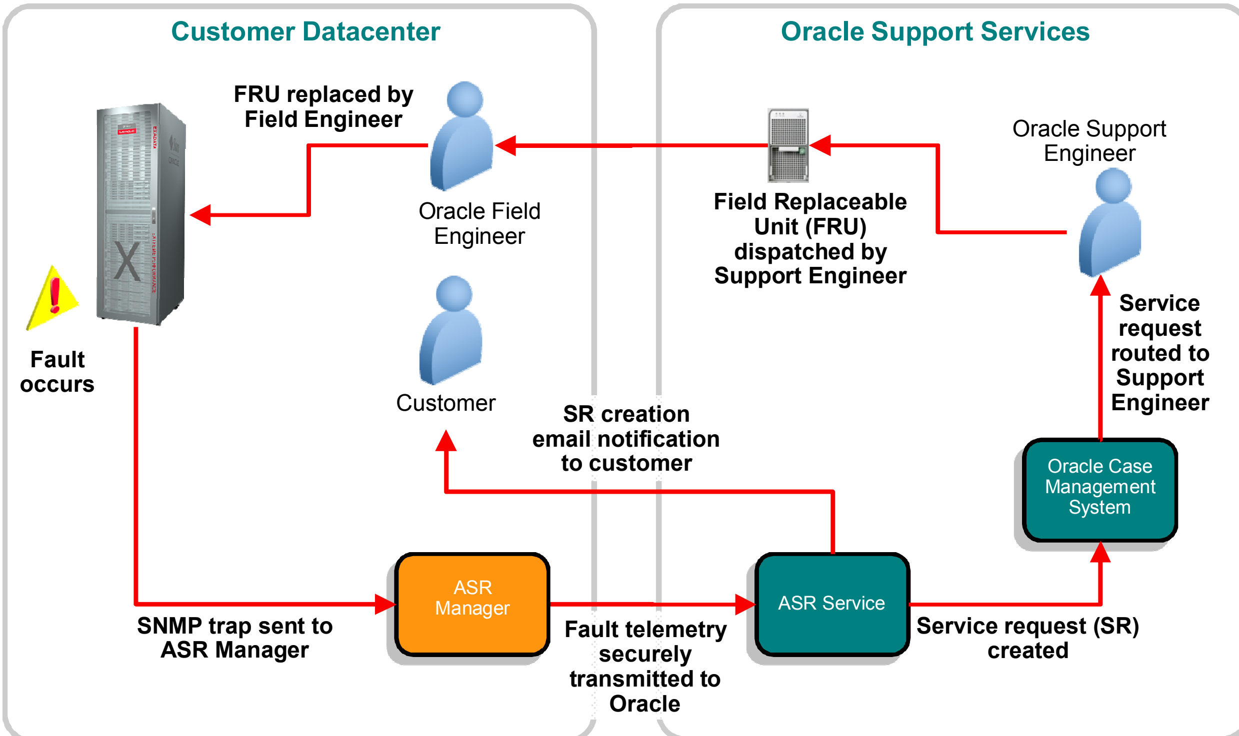
Auto Service Request: Overview

- Automatically and securely opens service requests with Oracle for common server hardware faults
  - Minimal data is collected and transmitted.
  - No IP addresses are communicated.
  - One-way (customer to Oracle) SSL encrypted communications are used.
- Provides fast and accurate resolution of hardware faults
  - Improved availability and less down time
- Can be integrated with existing monitoring tools
  - The ASR Manager can send SR notifications via SNMP traps to existing monitoring tools.
- Is included with hardware warranty and Oracle Premier Support for Systems

Auto Service Request (ASR) is a secure, scalable, and customer-installable software solution that is available as a feature of your Oracle hardware warranty and Oracle Premier Support for Systems. The ASR software helps to resolve problems faster by using auto-case generation for server and storage systems when specific hardware faults occur.

Since Exadata software version 11.2.1.3.1, ASR functionality is available for Exadata. ASR automatically opens service requests (SRs) with Oracle Support when specific hardware faults occur either in the Exadata Storage Servers or in the Exadata database servers. ASR covers the following server components: CPUs, disk controllers, disks, flash cards, flash modules, InfiniBand cards, memory modules, system boards, power supplies, and fans.

ASR Process



The diagram in the slide illustrates the ASR process. It shows how a service request (SR) is automatically opened by the ASR Manager after it receives an SNMP trap that is triggered by a server hardware fault. Within Oracle Support Services, the automatic service request is acknowledged via email and serviced according to normal procedures.

Customers must note that there are occasions when an SR may not be automatically filed. Foreexample, this can happen due to loss of connectivity to the ASR Manager. Customers can continue to monitor their systems for faults and engage Oracle Support if they notice a fault but do not receive notice that a service request has been automatically filed.

ASR Requirements

- A Linux or Solaris server capable of running the ASR Manager:
  - An existing ASR Manager (version 2.7 or greater) can be used to monitor Exadata.
  - It is possible, but not recommended to use an Exadata database server as the ASR Manager.
- Connectivity between the ASR Manager and the Exadata management network
- HTTPS connectivity (either directly or via a proxy) from the ASR Manager to Oracle Support
- My Oracle Support account with current contact information for all Exadata assets
- Exadata software version 11.2.1.3.1 or later on the database servers and Exadata Storage Servers

The slide lists the requirements for using ASR in conjunction with Exadata.

Information about the server requirements for the ASR Manager is available in the Hardware and Software Requirements section of the *Oracle Auto Service Request Exadata Database Machine Quick Installation Guide*.

The configuration of ASR can be performed as part of the initial installation and configuration of Exadata. Customers can opt out of configuring ASR. However, this is not recommended by Oracle.

The following section outlines the ASR configuration process, which can be used by existing Exadata customers to configure ASR after Exadata is already installed.

Configuring the ASR Manager

1. Install the ASR Manager components:
   – Oracle Automated Service Manager (OASM) Package
   – Service Tools Bundle (STB) for Solaris
   – Oracle Auto Service Request (ASR) Package
2. Register the ASR Manager:
   – As `root` on the ASM Manager server, run:

```
# asr register
```

   – Follow the prompts and provide the requested information.
   – To check the registration status, run:

```
# asr show_reg_status
```

   – To test the network connectivity, run:

```
# asr test_connection
```


Configuring a new ASR Manager server consists of two main tasks:

1. Installing the ASR Manager software components
2. Registering the ASR Manager

The procedure for installing the ASR Manager software components varies depending on the ASR Manager version and the platform selected to run the ASR Manager. Refer to the ASR documentation at http://www.oracle.com/technetwork/systems/asr/documentation/index.html for a detailed installation guide.

Configuring Exadata for ASR

1. Configure SNMP trap destinations:
   – As `root` on each database server, run:

```
DBMCLI> alter dbserver snmpsubscriber = -
>  ((host='<ASR Manager>', port=162, community=public, type=ASR, fromip=<eth0 IP>))
```

   – As `root` or `celladmin` on each Exadata Storage Server, run:

```
CellCLI> alter cell snmpsubscriber = -
>  ((host='<ASR Manager>', port=162, community=public, type=ASR))
```

   – Preserve existing SNMP registrations.

2. Verify ASR SNMP subscribers:
   – As `root` on each database server, run:

```
DBMCLI> list dbserver attributes snmpsubscriber
```

   – As `root` or `celladmin` on each Exadata Storage Server, run:

```
CellCLI> list cell attributes snmpsubscriber
```

ORACLE

Each Exadata database server and Exadata Storage Server must be configured to deliver SNMP traps to the ASR Manager server. Use the commands shown in the slide to configure and verify the required SNMP trap destinations. You may need to modify the suggested commands to preserve any existing SNMP configurations.

In the commands, any reference to `<ASR Manager>` must be replaced with the ASR Manager server host name or IP address. Also, `<eth0 IP>` must be replaced by the IP address for the database server eth0 network interface. The `fromip` argument is required only if you use Exadata Storage Server version 11.2.2.4.0 or later. Otherwise, use the same command, but do not include `fromip=<eth0 IP>`.

Remember that the SNMP subscriber configuration is required for every Exadata database server and Exadata Storage Server. You can use the `dcli` utility to perform the configuration simultaneously across multiple servers. For example, the following command can be used to perform the SNMP subscriber configuration on all the storage servers listed in the `cell_group` file:

```
# dcli -g cell_group -l celladmin -n "cellcli -e \
> alter cell snmpsubscriber=\(\(host=<ASR Manager>, \
> port=162, community=public, type=ASR\)\)"
```

Activating ASR Assets

- As `root` on the ASR Manager server:
  - Activate the ASR Manager host:

```
# asr activate_asset -i <ASR Manager IP>
```

  - Activate each Exadata database server ILOM and each Exadata Storage Server ILOM:

```
# asr activate_asset -i <Asset ILOM IP>
```

  - Activate each Exadata database server and each Exadata Storage Server:

```
# asr activate_exadata -i <Asset IP> \
>  -h <Asset hostname> -l <Asset ILOM IP>
```

  - Verify that all the nodes are visible on the ASR Manager:

```
# asr list_asset
```

- Activate ASR assets in My Oracle Support
  - Follow the procedure in My Oracle Support note 1329200.1.

ORACLE

The slide shows the commands required to activate the ASR assets. All commands are performed as `root` on the ASR Manager server. Activating the ASR Manager host is a one-time operation, whereas the other commands must be executed once for each Exadata database server and Exadata Storage Server.

In the commands:

- `<ASR Manager IP>` must be replaced with the ASR Manager server IP address
- `<Asset ILOM IP>` must be replaced with the database server or storage server ILOM interface IP address
- `<Asset IP>` must be replaced with the database server or storage server management network IP address
- `<Asset hostname>` must be replaced with the database server or storage server host name

After the Exadata server assets are activated in the ASR Manager, the servers are listed in My Oracle Support as Pending ASR Activations. To complete the activation process, the servers must be approved in My Oracle Support by using the process outlined in My Oracle Support note 1329200.1.

Verifying the ASR Configuration

- Verify the ASR configuration:
  - As `root` on each database server, run:

```
DBMCLI> alter dbserver validate snmp type=ASR
```

  - As `root` or `celladmin` on each Exadata Storage Server, run:

```
CellCLI> alter cell validate snmp type=ASR
```

ORACLE

After the assets are activated, the ASR configuration is completed and the system is enabled. To test the configuration and verify that it functions correctly, execute the commands shown in the slide. Each time a validation command is executed, the end-to-end ASR process is tested and an email is sent indicating that the ASR service inside Oracle Support has received the validation request. No service request is opened for validation requests.

Oracle Configuration Manager: Overview

Oracle Configuration Manager collects configuration information for propagation to Oracle Support.

- Benefits include:
  - Efficient problem diagnosis
  - Easier service request reporting
- Configuration information can be uploaded to Oracle either automatically or manually.
- Configuration information remains confidential.
- Oracle recommends the use of OCM in conjunction with Exadata.

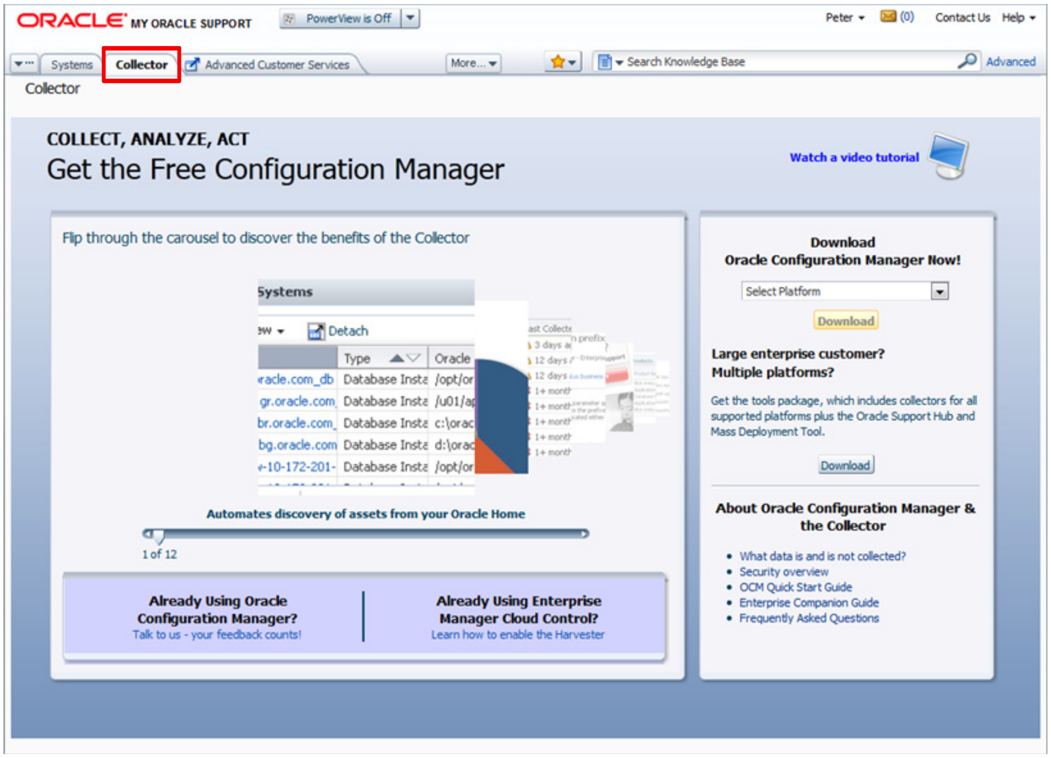
Oracle Configuration Manager (OCM) automatically collects configuration information from your environment at regular intervals. This configuration information can be uploaded to My Oracle Support. This helps Oracle maintain up-to-date information about your environment, diagnose support issues more efficiently, and offer consistently better support outcomes. OCM is recommended for use in conjunction with Exadata.

Oracle recommends that OCM is set up so that configuration information can be automatically collected and uploaded to the Oracle Customer Configuration Repository and My Oracle Support. Alternatively, information can be manually uploaded by an administrator.

No business or personal information is collected and uploaded, except for contact information, which is used for handling transmission problems. All information is kept strictly confidential.

Note that OCM is separate and distinct from the Enterprise Manager Configuration Management Pack.

Configuring Oracle Configuration Manager



ORACLE

Like ASR, configuration of OCM is typically performed as part of the initial installation and configuration of Exadata. Customers can opt out of configuring OCM. However, this is not recommended by Oracle.

For customers who want to configure OCM against an existing Exadata system, the easiest way is to use the resources on the *Collector* tab in My Oracle Support. The slide shows the interface presented to users, which they can use to get started with OCM.

The OCM collector is installed on the Exadata database servers. No OCM components are installed on the Exadata Storage Servers.

Quiz

Oracle recommends that the ASR Manager is configured on one of the Exadata servers.

a. True
b. False

Answer: b

The ASR Manager can be configured on an Exadata database server, but this is not recommended.

Quiz

Which of the following events causes ASR to create a service request with Oracle Support?

a. A disk fault in a storage server
b. An ORA-600 error in a database server
c. A power supply in a server fails
d. A hardware fault in a server fan
e. All of the above

Answer: a, c, d

Quiz

The Enterprise Manager Configuration Management Pack is required to use Oracle Configuration Manager.

a. True

b. False

Answer: b

Oracle Configuration Manager is separate and distinct from the Enterprise Manager Configuration Management Pack.

Summary

In this lesson, you should have learned how to:

- Describe the Auto Service Request (ASR) function and how it relates to Exadata
- Explain the implementation requirements for ASR
- Explain the ASR configuration process
- Describe Oracle Configuration Manager (OCM) and how it relates to Exadata

ORACLE

Additional Resources

- Documentation and recommendations:
  - Oracle Auto Service Request User Documentation and Product Qualification Information
    http://www.oracle.com/technetwork/systems/asr/documentation/index.html
  - Auto Service Request Quick Installation Guide for Oracle Exadata Database Machine
    http://docs.oracle.com/cd/E50790_01/doc/doc.121/e23333/toc.htm

24

Oracle Database Exadata Cloud Service Overview

ORACLE

Objectives

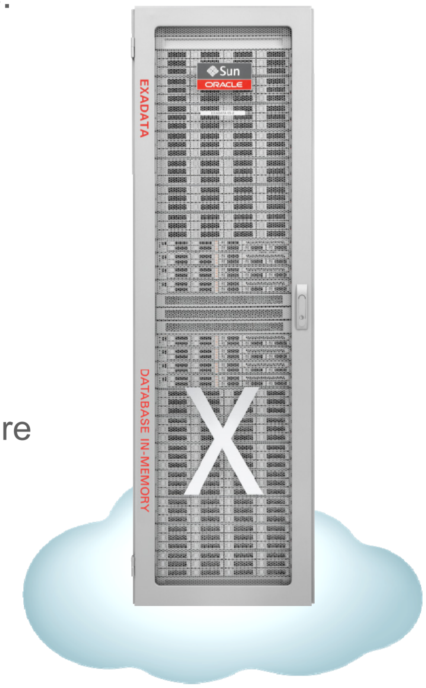
After completing this lesson, you should be able to:

- Describe the architecture and capabilities of Exadata Cloud Service
- Compare and contrast between Exadata Cloud Service and an on-premise Exadata implementation

ORACLE

Introducing Exadata Cloud Service

- Oracle Database with all features and options:
  - Industry-leading database for mission-critical OLTP and analytics
- On Exadata Database Machine:
  - The fastest and most available database cloud platform
- In the Oracle Cloud:
  - No capital expenditure, just a simple monthly subscription
  - Oracle deploys and manages the infrastructure
  - Fast, elastic, web-driven service provisioning
  - Complete service isolation with no over-provisioning
  - 100% compatibility with on-premises applications and Oracle database



ORACLE

Oracle Database Exadata Cloud Service enables you to leverage the power of Oracle Exadata Database Machine inside the Oracle Cloud. With Exadata Cloud Service you get:

- Oracle Database with all features and options:

Each Exadata Cloud Service database deployment is provisioned with a complete Oracle Database installation that includes all the features of Oracle Database Enterprise Edition, plus all the database enterprise management packs and all the Enterprise Edition options, such as Oracle Real Application Clusters (RAC), Oracle Database In-Memory, and Oracle Multitenant.

- On Exadata Database Machine:

The foundation for Exadata Cloud Service is Exadata Database Machine. Deployed at thousands of sites around the world, Exadata is established as the highest performing and most available platform for running Oracle Database. With a fault-tolerant architecture featuring scale-out database servers and scale-out intelligent storage connected with a fully redundant high-performance InfiniBand fabric, Exadata is an ideal cloud platform.

Exadata Cloud Service delivers all of the advanced features of Exadata, including SQL offload, Smart Flash Cache, Storage Index, and Hybrid Columnar Compression.

- In the Oracle Cloud:

Customers pay a simple monthly subscription fee for Exadata Cloud Service. There is no initial capital cost and no data center costs.

All supporting infrastructure for Exadata Cloud Service is deployed, maintained and managed by Oracle, including datacenter networking, private Exadata InfiniBand networks, physical Exadata database servers and storage servers, firmware, and Exadata Storage Server software.

Exadata Cloud Service includes easy-to-use web-based wizards through which you can quickly provision an Exadata system and associated database deployments. The wizards are available through Oracle's Cloud Portal, http://cloud.oracle.com.

Exadata Cloud Service features complete service isolation with no over-provisioning of hardware to ensure that response times and throughput are predictable for critical business processes. This contrasts with some cloud service delivery models that silently overprovision hardware, and consequently may not be able to deliver the expected resources during busy periods.

Exadata Service is 100% compatible with on-premises Oracle databases and all existing applications. Exadata Cloud Service enables existing on-premises Exadata customers to easily embark on a journey to the cloud – without compromising the database performance and availability levels they enjoy with their on-premises Exadata deployments. While existing Oracle Database customers who have not yet experienced Exadata can easily start enjoying the performance, availability and scalability benefits of Exadata without compromising any of the database functionality that they rely on. With Exadata Cloud Service, organizations can easily deploy a hybrid cloud environment that uses on-premises databases as well as databases in the Cloud.

Service Configuration Options

	Quarter Rack	Half Rack	Full Rack
Number of Database Servers	2	4	8
Number of CPU Cores [1]	16 - 68	56 - 136	112 – 272
Total RAM Capacity	496 GB	992 GB	1984 GB
Number of Exadata Storage Servers [2]	3	6	12
Total Flash Capacity	19.2 TB	38.4 TB	76.8 TB
Total Usable Disk Capacity [3]	42 TB	84 TB	168 TB
Maximum SQL Flash Bandwidth [4]	30 GB/sec	60GB/sec	120GB/sec
Maximum SQL Flash Read IOPS [5]	900,000	1,800,000	3,600,000
Maximum SQL Flash Write IOPS [6]	500,000	1,000,000	2,000,000
Maximum SQL Disk Bandwidth [4]	4.5 GB/sec	9 GB/sec	20 GB/sec
Maximum SQL Disk IOPS [5]	7,000	14,000	28,000
Maximum Data Load Rate [7]	5 TB/hour	10 TB/hour	10 TB/hour

ORACLE

Each Exadata Cloud Service instance is based on an Exadata configuration that contains a predefined number of database servers and a predefined number of Exadata Storage Servers, all tied together by a high-speed, low-latency InfiniBand network and intelligent Exadata software. Three configurations are offered; Quarter Rack, Half Rack and Full Rack. The table in the slide outlines the vital statistics for each system configuration.

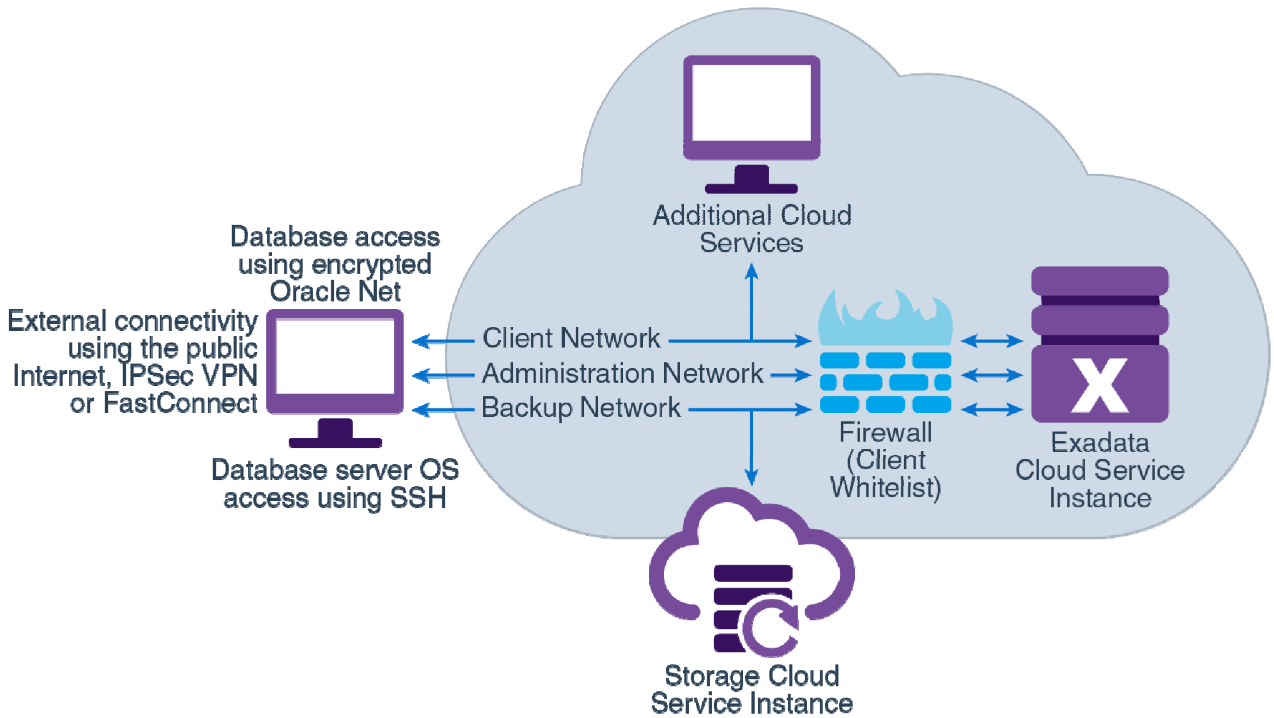
Note the following:

1. Each Exadata Cloud Service configuration is equipped with a fixed amount of memory, storage and network resources. However, you can choose how many database server CPU cores are enabled, within the minimum and maximum limits listed in the table. This enables you to scale an Exadata Cloud Service configuration to meet workload demands, and only pay for the database server resources that you need. Each database server must contain the same number of enabled CPU cores.
2. Note that Exadata Cloud Service Half Rack and Full Rack configurations differ from on-premises Half Rack and Full Rack. The Exadata Cloud Service Half Rack contains six storage servers, and the Exadata Cloud Service Full Rack contains twelve storage servers. Consequently, an Exadata Cloud Service Half Rack is twice the size of a Quarter Rack, and a Full Rack is twice the size of a Half Rack.
3. The usable storage capacity is the storage that available for Oracle Database files after

taking into account high-redundancy ASM mirroring (triple mirroring), which is used to provide highly resilient database storage on all Exadata Cloud Service configurations. The usable storage capacity does not factor in the effects of Exadata compression capabilities, which can be used to increase the effective storage capacity.

4. Bandwidth is the peak physical scan bandwidth that is achieved by running SQL, and assuming that there is no database compression. Effective user data bandwidth is higher when database compression is used. In all cases, actual performance varies by application.
5. Based on 8K Oracle Database I/O requests. Note that the I/O size greatly affects Flash IOPS, so IOPS based on smaller I/Os is not relevant for databases.
6. Based on 8K Oracle Database I/O requests. Flash write I/Os are measured at the storage servers after ASM mirroring, which issues multiple storage I/Os to maintain redundancy.
7. Load rates are typically limited by database server CPU capacity, not I/O. Rates vary based on a variety of factors, including the load method and data types used, and the use of indexes, compression and partitioning.

Service Connection Options



The diagram in the slide outlines the primary ways of connecting to Exadata Cloud Service.

You can connect directly from cloud clients, which are applications already running in the Oracle Cloud. This includes java applications connecting through JDBC running on the Java Cloud Service, or client applications connecting through Oracle Net (SQL*Net) running on a Compute Cloud Service instance.

You can also connect to Exadata Cloud Service from your on-premises applications by using Oracle Net. On Exadata Cloud Service, Oracle Net is configured by default to use native encryption and integrity capabilities to secure data in transit.

Access to the database server operating system is provided using Secure Shell (SSH). This is primarily used for administration purposes.

A backup network is also provided. This network keeps high-load activities separate from application connections and is primarily used when Exadata Cloud Service database deployments are backed up to an Oracle Storage Cloud Service container.

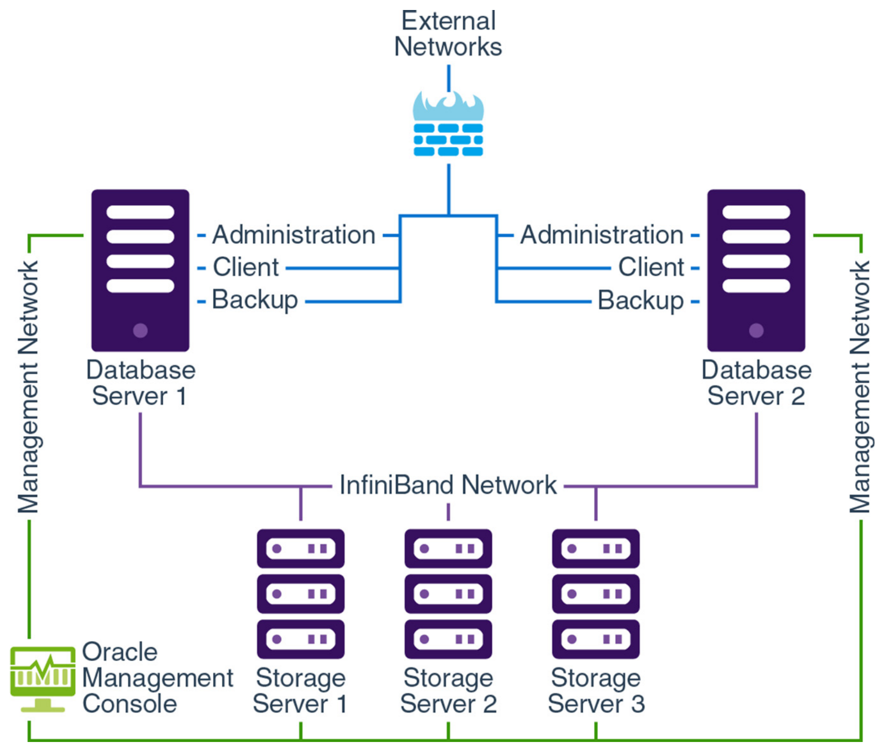
The Oracle Cloud network is firewall-protected and you must specify the clients that can access the service. The permitted clients are registered in the firewall using a whitelist.

To provide connectivity between your network and the Oracle Cloud you can:

- Use the public Internet and secure protocols, such as SSH and encrypted Oracle Net.

- Configure an IPSec VPN to enable secure connectivity between your network and the Oracle Public Cloud over the Internet.
- Use Oracle Network Cloud Service - FastConnect Partner Edition, which is a secure, high-bandwidth, point-to-point networking solution that is integrated with Oracle Cloud.

Service Architecture



The diagram in the slide illustrates the Exadata Cloud Service architecture for a Quarter Rack service instance. Larger service instances are principally the same, except that they contain more database servers and Exadata Storage Servers.

The architecture is essentially the same as for an on-premises implementation of Exadata, with clustered database servers connected to a grid of Exadata Storage Servers through a high-speed, low-latency, InfiniBand network.

Application users and administrators can connect only to the database servers, using the supplied administration, client and backup network interfaces.

Oracle manages all hardware, firmware, and the Exadata Storage Server software by using a separate management network.

Service Availability

- Exadata Cloud Service inherent HA capabilities:
  - Exadata platform full hardware redundancy:
    - Database and storage servers
    - InfiniBand and Ethernet networking
    - Power supplies and PDUs
  - Oracle software HA capabilities:
    - RAC protects against database server failures
    - ASM data mirroring protects against storage failures
      - Exadata Cloud Service used 3-way mirroring (ASM high redundancy)
    - Plus Flashback technologies, Online DDL, In-memory fault-tolerance, and so on...
- Implementation best-practices are derived from thousands of mission-critical deployments worldwide

ORACLE

Exadata Cloud Service inherits all of the high availability (HA) capabilities that are an integral part of the Exadata platform. Exadata has full hardware redundancy including redundant InfiniBand networking, redundant network ports, redundant Power Distribution Units (PDUs), redundant power supplies, redundant database servers, and redundant storage servers.

Exadata Cloud Service inherits all of the Oracle Database HA capabilities. Oracle RAC protects against database server failures. Automatic Storage Management (ASM) provides data mirroring to protect against disk or storage server failures. In Exadata Cloud Service, high redundancy (3-way mirroring) is used by default to protect data. Oracle Recovery Manager (RMAN) provides extremely fast and efficient backups to disk or cloud storage. Oracle Database Flashback technologies allow user errors to be backed out at the database, table or row level. Online Data Definition Language (DDL) operations enhance availability during data maintenance operations. In-memory fault-tolerance mirrors the contents of the in-memory data store, which enables surviving stores to automatically and transparently satisfy user queries if a database instance fails.

Because of its HA pedigree, Exadata is deployed worldwide at thousands of mission-critical deployments in leading banks, airlines, telecommunications companies, stock exchanges, government agencies and e-commerce sites. The best practices guiding the implementation of Exadata Cloud Service are derived from the experience of those deployments.

Service Scalability

Exadata Cloud Service instances support:

- Scaling within an Exadata system:
  - Adjust the number of database server CPU cores

	Quarter Rack	Half Rack	Full Rack
Minimum number of enabled CPU cores	16 total (8 per server)	56 total (14 per server)	112 total (14 per server)
Maximum number of enabled CPU cores	68 total (34 per server)	136 total (34 per server)	272 total (34 per server)

  - There are no changes to other resources
- Scaling between Exadata systems
  - Move a database to a different system
    - For example: Quarter Rack to Half Rack
  - Adjust processing power, storage capacity, number of database servers, and number of storage servers

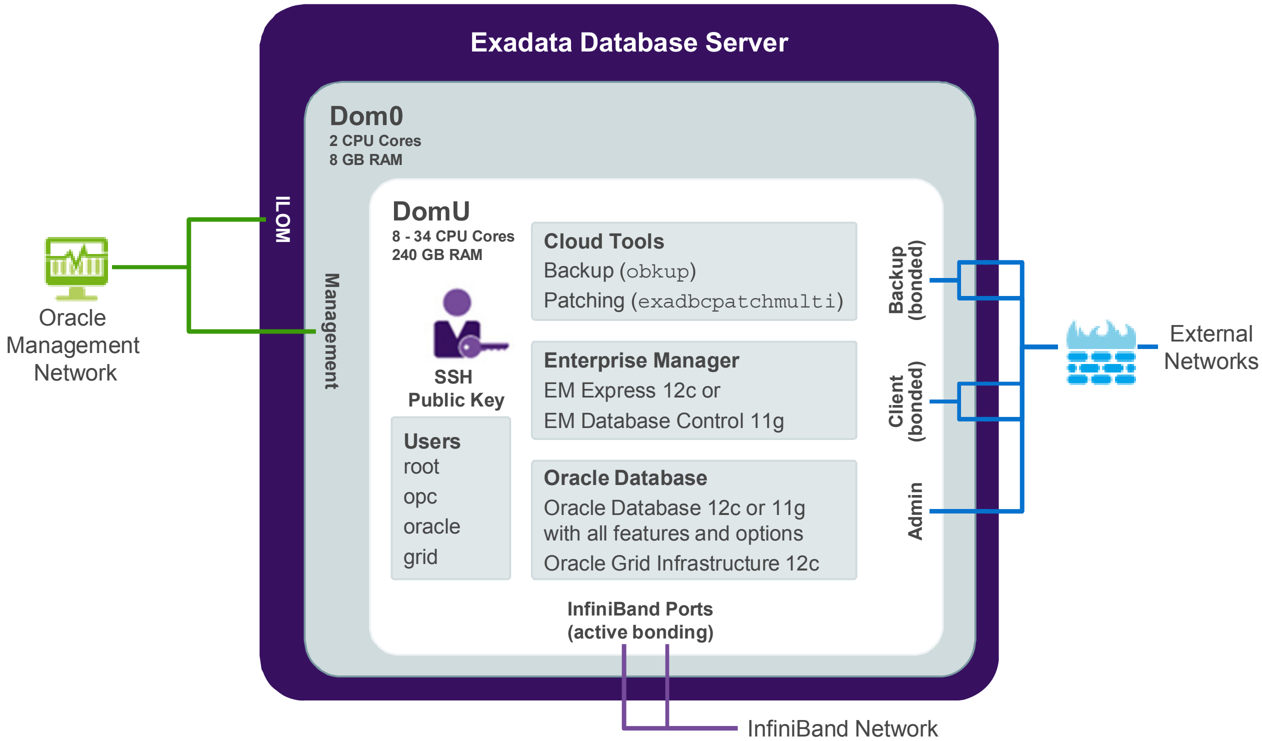
ORACLE

Oracle Database Exadata Cloud Service instances support two kinds of scaling:

- Scaling within an Exadata system enables you to add or remove database server processing power within the confines of your existing Exadata system. The minimum and maximum number of enabled CPU cores depends on the system configuration being used, as well as any additional limits associated with your subscription. When you scale within an Exadata system all other resources remain the same, such as the storage capacity, the number of database and the number of storage servers.
- Scaling between Exadata systems enables you to move a database to a different Exadata system configuration. For example, from a Quarter Rack to a Half Rack. This is useful when database deployments require:
  - Processing power that is outside the scope of the current system configuration.
  - Storage capacity that is outside the scope of the current system configuration.
  - A performance boost that can be delivered by increasing the number of available compute nodes.
  - A performance boost that can be delivered by increasing the number of available Exadata Storage Servers.

  Scaling across systems requires that the data associated with your database deployments are backed up and restored on a different Exadata Database Machine.

Service Access and Security



The diagram in the slide drills into more detail for each Exadata database server and outlines key architectural components.

Each Exadata database server contains one Virtual Machine (VM), known as DomU, running on a VM hypervisor, known as Dom0. This configuration ensures a distinct separation between the Oracle-managed and customer-managed components.

Oracle manages the hardware, firmware and Dom0 by using the Integrated Lights Out Manager (ILOM) and Dom0 management network interfaces. Customers have no access to these interfaces.

Customers have secure access to DomU using the administration, client and backup networks. The backup and client networks use bonded network interfaces to maximize performance and availability.

By default, network access to DomU is limited to SSH connections on port 22. To access network protocols and services on DomU by using a port other than port 22 customers must specifically configure network access. This is also required to enable database access by using Oracle Net. Options include:

- Enable access to a specific network service. Customers must raise a service request with Oracle to reconfigure the network firewall and enable designated clients to access specific network ports.

- Use an SSH tunnel. An SSH tunnel enables you to access a specific network service by using an SSH connection as the transport mechanism.
- Configure an IPSec VPN. This enables secure connectivity between a customer network and the Oracle Public Cloud over the Internet.
- Use Oracle Network Cloud Service - FastConnect Partner Edition, which is a secure, high-bandwidth, point-to-point networking solution that is integrated with Oracle Cloud.

Operating system security for DomU is based on an SSH public/private key pair. During configuration for Exadata Cloud Service, customers register a public key. Thereafter, the corresponding private key must be provided in order to connect to DomU using SSH. At all times, the customer retains the private key that enables access to the DomU operating system.

As a result of this configuration, customers manage DomU and all of the Oracle software that it contains. Oracle provides assistance, through the provision of cloud tools that simplify backup and patching operations, but ultimately it is the responsibility of the customer to perform these operations.

Data Security

In addition to the database server security provisions, Oracle Database data is secured:

- At Rest
  - Tablespaces are encrypted using Transparent Data Encryption
  - Data cannot be viewed or tampered with by accessing the Exadata Storage Servers
- In Transit
  - Client application connections are encrypted using native Oracle Net encryption and integrity capabilities
  - Data cannot be viewed or tampered with as it passes over the network

ORACLE

In addition to the Database Server security measures, data security in Exadata Cloud Service is enabled by default in two ways:

- Security of Data at Rest

  Oracle databases created on Exadata Cloud Service automatically use the Transparent Data Encryption (TDE) feature for tablespace encryption. This ensures that data cannot be viewed or tampered with by accessing the Exadata Storage Servers.

- Security of Data in Transit

  Oracle databases created on Exadata Cloud Service automatically use native Oracle Net encryption and integrity capabilities. Encryption of network data provides data privacy so that unauthorized parties are not able to view data as it passes over the network. In addition, integrity algorithms protect against data modification and illegitimate replay.

All of the configuration settings for TDE and native Oracle Net security are maintained in the Exadata database servers, which are secured by using the mechanisms described earlier in the lesson.

Management Responsibilities

Oracle Managed (No Customer Access)	Customer Managed (No Oracle Access)
Initial configuration and installation	Database server DomU - including OS
Exadata Storage Server software and objects	Oracle Database, Grid Infrastructure, ASM
Database server Dom0	Database and OS updates*
InfiniBand switches, HCAs, and partitioning	Database and OS monitoring*
Management and ILOM networks	Database backup and recovery*
All hardware, firmware, and BIOS	
Client access VLANS and IP addresses	* Oracle tooling provided

ORACLE

The table in the slide summarizes the division of management responsibilities for Exadata Cloud Service. Additional details are provides in the following notes and throughout this lesson.

Oracle manages the following elements, which customers cannot access:

- Initial configuration and installation of all hardware and software components.
- Exadata Storage Server software and all Exadata Storage Server objects (cell disks, grid disks, flash cache and so on).
- Database server Dom0 environment, including the operating system (OS) and hypervisor.
- InfiniBand switches, Host Channel Adapters (HCAs), and InfiniBand partitioning.
- Dom0 management network and Integrated Lights Out Manager (ILOM) network.
- All hardware, firmware and BIOS.
- IP addresses and VLAN configuration for the client and backup networks.

Customers manage the follow elements, which Oracle cannot access:

- Database server virtual machine (DomU), including the kernel and other OS components.
- Oracle Database, Grid Infrastructure, and Automatic Storage Management (ASM). Customers manage all aspects of their Oracle databaseroless, including creation and maintenance of tablespaces, tables, indexes, users,, and so on.

- Patching and updates for Oracle Database software and DomU OS components. Oracle provides Exadata Cloud Service-specific tools to assist.
- Oracle Database and OS monitoring. Exadata Cloud Service is configured with Enterprise Manager Database Control 11*g* and Enterprise Manager Express 12*c* for local database management. Customers may manage their Exadata Cloud Service environment using a separate Oracle Enterprise Manager Cloud Control environment.
- Database backup and recovery. Oracle provides Exadata Cloud Service-specific tools to assist. Also, during database configuration you can choose to configure automated daily database backups.

Storage Configuration

- Preconfigured ASM Disk Groups:
  - High redundancy disk groups, which consume nearly all of the storage
    - DATA contains Oracle Database data files
    - RECO contains the Fast Recovery Area (FRA)
  - System disk groups, which are comparatively very small
    - DBFS contains shared clusterware files
    - ACFS disk groups contain Oracle Database binaries and patch files
- Configuration options for space allocation:
  - Provision for Local Backups: 40% DATA, 60% RECO
  - No Provision for Local Backups: 80% DATA, 20% RECO

	Quarter Rack	Half Rack	Full Rack
Total Usable Disk Capacity	42 TB	84 TB	168 TB
Disk Group Sizes ***with*** Provision for Local Backups	DATA: 16.8 TB RECO: 25.2 TB	DATA: 33.6 TB RECO: 50.4 TB	DATA: 67.2 TB RECO: 100.8 TB
Disk Group Sizes ***without*** Provision for Local Backups	DATA: 33.6 TB RECO: 8.4 TB	DATA: 67.2 TB RECO: 16.8 TB	DATA: 134.4 TB RECO: 33.6 TB

ORACLE

As part of provisioning each Oracle Exadata Database Machine environment, the storage space inside the Exadata Storage Servers is provisioned for use by Oracle Automatic Storage Management (ASM). By default, the following ASM disk groups are created:

- Two high redundancy disk groups consume nearly all of the available storage space:
  - The DATA disk group is intended for the storage of Oracle Database data files.
  - The RECO disk group is primarily used for storing the Fast Recovery Area (FRA), which is an area of storage where Oracle Database can create and manage various files related to backup and recovery, such as RMAN backups and archived redo log files.
- The system disk groups support various operational purposes.
  - The DBFS disk group is primarily used to store the shared clusterware files (Oracle Cluster Registry and voting disks).
  - The ACFS disk groups are primarily used to store Oracle Database binaries and to facilitate patching.

  Compared to the DATA and RECO disk groups, the system disk groups are very small, consuming less than 1% of the total available disk space. The system disk groups are configured to use normal redundancy. You should not store Oracle Database data files or backups inside the system disk groups.

Note that the disk group names contain a short identifier string that is associated with your Exadata Database Machine environment. For example, the identifier could be `C2`, in which case the DATA disk group would be named `DATAC2`, the RECO disk group would be named `RECOC2`, and so on.

As an input to the provisioning process, you must decide if you intend to perform backups to the local storage within your Exadata Database Machine. Your backup storage choice profoundly affects how storage space in the Exadata Storage Servers is allocated to the ASM disk groups.

If you choose to provision for local backups, approximately 40% of the available storage space is allocated to the DATA disk group and approximately 60% is allocated to the RECO disk group. If you choose not to provision for local backups, approximately 80% of the available storage space is allocated to the DATA disk group and approximately 20% is allocated to the RECO disk group.

After the Exadata Database Machine is activated, the only way to adjust the storage allocation is by lodging a Service Request with Oracle. For details see My Oracle Support Note 2007530.1.

The table in the slide outlines how the usable storage capacity is allocated to the DATA and RECO disk groups for each configuration option. The usable storage capacity is the storage that available for Oracle Database files after taking into account high-redundancy ASM mirroring (triple mirroring), which is used to provide highly resilient database storage on all Exadata Cloud Service configurations. The usable storage capacity does not factor in the effects of Exadata compression capabilities, which can be used to increase the effective storage capacity. The usable disk capacity does not include the space allocated to the system disk groups.

Storage Management Details

- Storage is preconfigured and allocated to ASM disk groups
- Customers manage database objects inside ASM
- Oracle manages the Exadata storage, including Exadata Storage Server software updates
- Exadata Storage Servers are configured using best practices:
  - One cell disk on each physical storage device
  - One set of grid disks for each ASM disk group
    - Space allocation depends on backup configuration
    - Disk groups for database files use high redundancy
    - Grid disk names are prefixed with the corresponding disk group name
  - Flash cache and flash log are preconfigured
    - Flash cache mode is write-through
  - IORM plan is active, IORM objective is set to balanced
- Customers have no direct access to the Exadata Storage Servers
  - Can request custom configurations for grid disks and IORM

ORACLE

With Exadata Cloud Service, storage is preconfigured and allocated to ASM disk group, so customers must manage database objects inside ASM. This includes management of tablespaces and the data they contain. Customers must also monitor and maintain ASM to ensure continued availability.

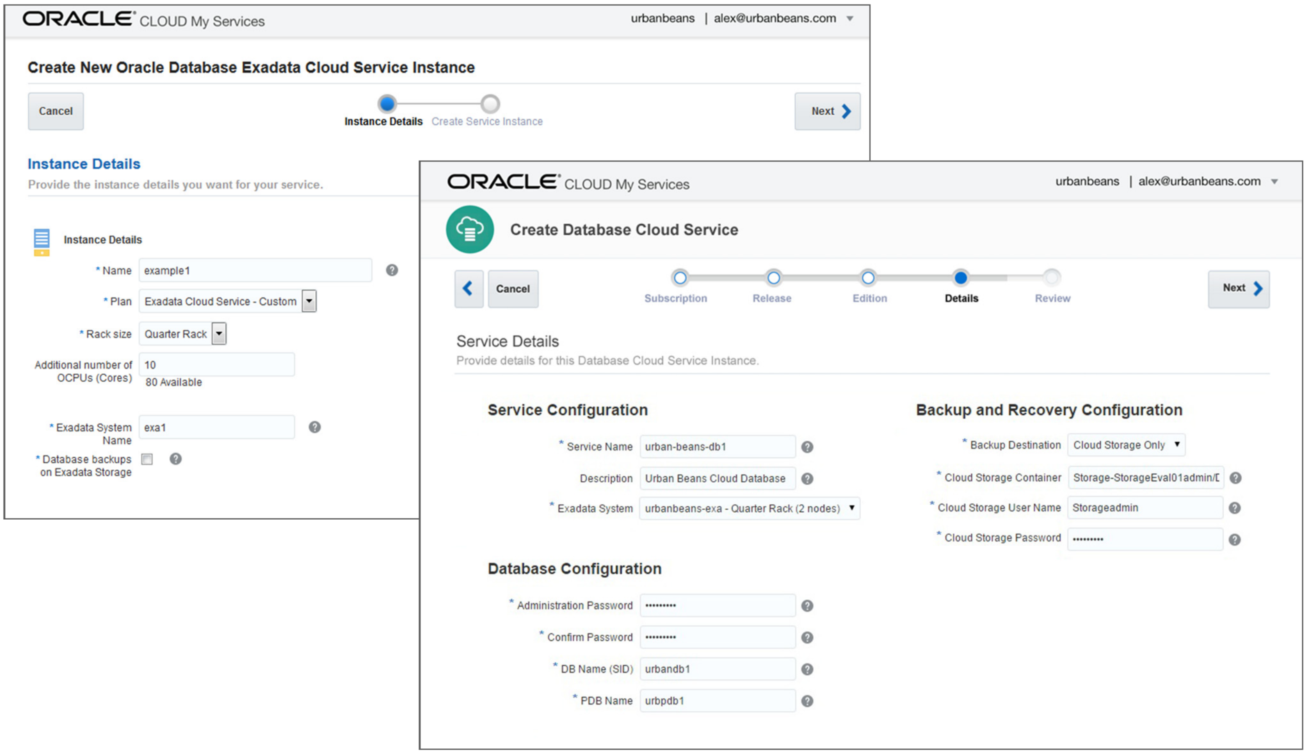
Oracle manages the Exadata Storage Servers, including Exadata Storage Server software updates. The Exadata Storage Servers are configured in accordance with best practices:

- The cell disks, and grid disks to support the DATA, RECO and system disk groups are preconfigured and use standard Exadata default settings.
- The flash cache and flash log are preconfigured. For Exadata Cloud Service, the flash cache is configured in write-back mode by default.
- The I/O Resource Management (IORM) plan is active and the objective is set to balanced.

No direct access is provided to the Exadata Storage Servers. However, you can lodge a Service Request (SR) for:

- Storage reconfiguration to redistribute the space amongst the existing grid disks, or to create additional grid disks.
- IORM configuration to a custom specification.

Simple Web-Based Provisioning

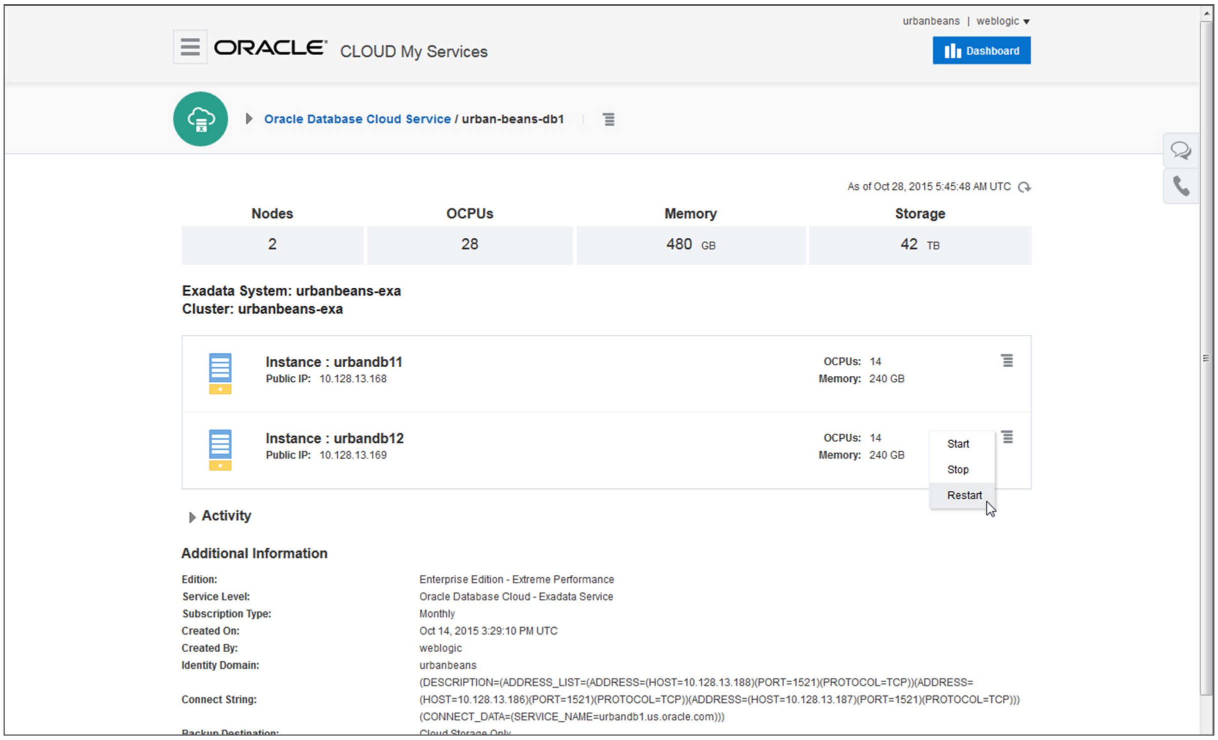


ORACLE

Exadata Cloud Service provides simple web-based provisioning interfaces, which you can use to quickly and easily provision an Exadata system and subsequently create their Oracle databases.

The screenshots in the slide show examples of the key screens in the Exadata Cloud Service instance provisioning wizard, and the wizard used to create databases on Exadata Cloud Service. In both cases, only a few key inputs are required and either task can be completed in only a few minutes.

Simple Web-Based Management



Exadata Cloud Service also provides a set of web-based management interfaces, which you can use to quickly and easily accomplish various management tasks including:

- Database server VM lifecycle activities (start, stop, restart).
- SSH key management (add, update).
- View service instance details.
- View database deployment details.
- Configure an Exadata I/O Resource Management (IORM) inter-database plan.
- Scale within an Exadata system (add or remove database server CPU cores).

The screenshot in the slide show a simple example of the interface that displays database deployment details along with the options to start, stop, or restart one of the database server virtual machines.

REST APIs

- REST APIs provide programmatic management and control:
  - Create or Delete a Database Deployment
  - Stop, Start or Restart a Database Server VM
  - View Details
    - View a Database Deployment
    - View All Database Deployments
    - View Database Servers
    - View the Status of an Operation

```
$ curl --include --request POST --cacert ~/cacert.pem
--user serviceadmin:Pa55_word
--header "X-ID-TENANT-NAME:usexample”
--header "Content-Type:application/json"
--data '{ "lifecycleState" : "stop", "vmName" : "node02" }‘
https://dbaas.oraclecloud.com/paas/service/dbcs/api/v1.1/instances
/usexample/db12c
```

ORACLE

In addition to the web-management interfaces, Exadata Cloud Service provides a series of REST APIs. The Exadata Cloud Service REST APIs enable you to perform various management operations programmatically, such as create and delete database instances, start and stop compute nodes, and view status information. The slide shows an example that uses the cURL utility to invoke the REST API call to stop a database server virtual machine.

See http://docs.oracle.com/cloud/latest/exadatacs/EXARS/index.html for details about the Exadata Cloud Service REST APIs.

Backup and Recovery

- Exadata Cloud Service provides automatic database backups:
  - Based on RMAN and exposed through the `obkup` utility
  - Available options:
    - Automatic backup to both cloud storage and local storage
      - Two separate backups
      - Only possible on systems provisioned for local backups
    - Automatic back to cloud storage only
    - No automatic backups
  - Default settings for automatic backups:
    - Scheduled daily at 1:01 AM
    - 7 day backup cycle with one full (RMAN level 0) backup and daily incremental (RMAN level 1) backups
    - 30 day retention period for cloud backups, 7 days for local backups
    - Oldest incremental backup is merging into the oldest full backup
    - Backup data is automatically encrypted
- User can take on-demand database backups using `obkup` or RMAN
- Oracle automatically backs up the database server DomU
  - Provides recovery from database server corruption

ORACLE

Exadata Cloud Service provides a backup feature that backs up the Oracle database associated with the database deployment. This feature is built over Oracle Recovery Manager (RMAN), and exposed through a simple utility command (`obkup`) that is installed on your Exadata system. It also relies on Oracle Database Backup Cloud Service, which in turn uses an Oracle Storage Cloud Service container for cloud storage.

When you create a database deployment on Exadata Cloud Service, you must choose from the following backup configuration options:

- **Both Cloud Storage and Local Storage**. Automatic backups and stored both on local Exadata storage and on an Oracle Storage Cloud Service container. This option is only available if you provisioned for local backups in your Exadata Cloud Service instance.
- **Cloud Storage Only**. Automatic backups are configured and stored only on an Oracle Storage Cloud Service container.
- **None**. No automatic backups are configured.

Following is an outline of the default backup configuration settings:

- Automatic backups are scheduled daily at 1:01 AM.
- Backups consist of a full (RMAN level 0) backup of the database, followed by daily incremental (RMAN level 1) backups, with a 7 day cycle between full backups.

- The retention period defines the period for which backups are maintained, as follows:
  - **Both Cloud Storage and Local Storage:** This option provides for two separate backups with different retention periods. The backup to local storage has a 7 day retention period and the backup to cloud storage has a 30 day retention period.
  - **Cloud Storage Only:** 30 days.
- After the initial retention period, each day when an incremental backup is taken, the oldest daily incremental backup is automatically merging into the oldest full backup.
- The backup data is automatically encrypted.

In addition to the automatic backup schedule, you can take on-demand database backups. You can run `obkup` directly to create an on-demand backup that follows the current automatic database backup configuration. Or you can perform custom backup and recovery operations using RMAN directly.

Apart from database backup facilities, Oracle automatically takes a periodic backup of each database server virtual machine (DomU). Oracle maintains there backups and will restore a backup in response to a catastrophic event, such as a corruption that disables the DomU environment.

Patching Exadata Cloud Service

- Customers apply database server patches (DomU)
  - Oracle provides quarterly patch set updates
    - Updates contain all recommended patches
  - Oracle provides `exadbcpatchmulti` to simplify patching:
    - Checks prerequisites
    - Applies patches
    - Lists applied patches
    - Performs patch rollback
  - Patches are applied in a rolling manner wherever possible
- Oracle patches all other components
  - Patches scheduled quarterly
  - Patch schedule notified in advance
  - Patches are applied in a rolling manner wherever possible

ORACLE

With Exadata Cloud Server, it is the customers responsibility to apply database server patches (to DomU). Oracle provides assistance by publishing quarterly patch set updates (PSUs), which contain all of the recommended patches. Also, Exadata Cloud Service contains the `exadbcpatchmulti` utility, which simplifies patching by coordinating patching tasks across all of the database servers. Apart from applying patches, the `exadbcpatchmulti` utility can perform prerequisite checks, list applied patches and even rollback patches. Wherever possible, patches are applied in a rolling manner, one server at a time, in order to minimize service downtime.

Oracle patches all other components, including the Exadata Storage Server software, the database server Dom0 environment, and all firmware. Patches are scheduled quarterly and the details are notified to customers in advance. Wherever possible, patches are applied in a rolling manner, one component at a time, in order to minimize service downtime.

Migrating to Exadata Cloud Service

- Logical Migration options include:
  - Oracle Data Pump
  - Oracle GoldenGate
- Physical Migration options include:
  - Recovery from an RMAN backup
  - Transportable Tablespaces
  - Oracle Data Guard

Full compatibility between existing on-premises Oracle databases and Oracle databases on Exadata Cloud Service makes migration easy and low risk. Following established best practices for Oracle Database, two types of migration methodologies are supported:

- Logical Migration allows data reorganization as part of migration. Database solutions that can be used for this purpose include Oracle Data Pump and Oracle GoldenGate.
- Physical Migration, which involves a byte-to-byte copy of the data, offers the simplest way to migrate databases. Solutions that can be used for this purpose include RMAN backup, Transportable Tablespaces, and Oracle Data Guard. You can also restore from a backup stored on the Oracle Public Cloud through the Oracle Database Backup Service.

Summary

In this lesson, you should have learned how to:

- Describe the architecture and capabilities of Exadata Cloud Service
- Compare and contrast between Exadata Cloud Service and an on-premise Exadata implementation