Oracle Maximum
Availability Architecture

Installing Oracle Extended Clusters on Exadata Database Machine

Best Practices for Oracle Database 12*c* Release 2

ORACLE WHITE PAPER | JUNE 2017

ORACLE

.

Introduction

Oracle Exadata Database Machine inherently has many High Availability (HA) capabilities providing redundancy and data protection for Oracle databases[1]. With Oracle Grid Infrastructure 12c release 2 (12.2), Oracle Grid Infrastructure supports the option of configuring cluster nodes in different locations as an Oracle Extended Cluster, which consists of nodes that are located in multiple sites.

By deploying Oracle 12.2 Extended Clusters on Exadata, you can expand and compliment HA benefits by providing availability for a localized site failure. This is particularly beneficial when there are isolated sites or availability domains (sometimes referred to as "fire cells" with independent power, cooling and resources) within a data center or between two metro data centers. With a properly configured Extended Cluster on Exadata, applications and databases can tolerate a complete site failure, and also any additional Exadata storage cell or Exadata database server failure. To reap the full Exadata MAA benefits including Disaster Recovery, Oracle recommends adding a standby Exadata rack in a remote location. This provides a comprehensive data protection, as well as fast failover of the RAC cluster in case of site failure.

This document will cover how to install and configure Oracle 12.2 Extended Clusters on Exadata. It is assumed that the reader has experience with installing and administering Oracle Real Application Clusters and is familiar with basic Exadata installation and configuration best practices.

Prerequisites

The following prerequisites are required:

- At least three physically separate sites (or availability domains) must be available.
  - Two separate sites will have Exadata DB servers and Exadata storage servers (cells) installed. The two sites must be on the same IB fabric and within 100m of an InfiniBand (IB) switch as follows:
    - If the total IB cable distance between sites is less than 100m, then cabling will extend between racks.
    - If the total IB cable distance between sites is greater than 100m but less than 200m, then a pair of spine switches must be placed between the two sites and within 100m of each site to permit a total distance of 200m.
    - See the Oracle Exadata Database Machine Extending and Multi-Rack Cabling Guide for details on connecting the leaf switches to in-between spine switches. The topology is the same as a 2-rack interconnection with spine switches in each rack.
  - Each site must consist of at least two DB servers and three cells
  - A completely separate quorum disk site that is accessible via NFS using the client Ethernet interface; this site must be in a separate site from a middle IB spine switch (if used to extend the cluster to 200m)
- A third, Quorum site must have the following configuration:
  - NFS server on OEL 6
  - Preferably two - 10G or 1G Ethernet interfaces bonded together (reliable low-latency network)
  - Typically, less than 50Km from each site

[1] http://www.oracle.com/technetwork/database/features/availability/exadata-maa-131903.pdf

- At least 50GB free space available to export via NFS

- Oracle Grid Infrastructure (GI) and Database (DB) version 12.2.0.1 or higher (DB versions earlier than 12.2 are not supported)
- Exadata software 12.2.1.1.1 or higher
- Latest version of OEDA able to install 12.2.0.1 Grid Infrastructure and DB software
- Sufficient space for database objects and redundancy
- EXTENDED REDUNDANCY diskgroups require four times the space used by database files *plus* an additional 15% of the total diskgroup size as free space to allow successful rebalance after the loss of a single disk. For example, if database files use 10 TB of space, then you should ensure that there is at least 40 TB for all ASM extent copies plus an additional 6 TB to rebalance the loss of a disk for a total of 46 TB. For this example, there should always be a minimum of 6 TB of free space in the diskgroup to permit rebalancing the loss of a single grid disk. Extended redundancy diskgroups will not be dismounted unless all partner disks or cells on one site are lost plus an additional disk or cell on the surviving site.

- Please see Requirements for Configuring RAC Extended Clusters on Exadata (Doc ID 2234289.1) for patches needed.

Step-by-Step Installation

These are high level steps involved to setup extended cluster for Exadata:

1. Install NFS server at a completely independent site
2. Create OEDA Installation XML File
3. Install and configure Exadata racks in desired sites
4. Use OEDA for preliminary configuration
5. Manually configure NFS server and client
6. Manually Configure Exadata Cells
7. Manually install Grid Infrastructure
8. Ensure ASM is using Best Practice Settings
9. Manually Create a RECO Diskgroup
10. Install Oracle Database Software
11. Install database using the Exadata Database Assistant (EDA)
12. Validate configuration

The details below assume the Exadata extended cluster will consist of two sites plus a quorum site. Each of the two cluster sites will consist of two DB servers and three cells. The following sections show the detailed steps required to perform the installation and configuration for the extended clusters.

Before starting the installation, be sure to choose the name for each of the sites. This name will be used in the cell and the ASM failure group configurations. The name should be reasonably short (the limit is 15 characters on the cells).

Install NFS server at a completely independent site

The NFS server will be configured as a quorum failure group to host the cluster voting files. Install the server with the following:

- OL 6 with NFS server installed
- Two 10g Ethernet interfaces bonded together and visible on the DB servers' client network (i.e.,bondeth0 on the DB servers)

» The server should be completely independent of the two other cluster sites in terms of power, HVAC, etc.

Create OEDA Installation XML File

Use the latest OEDA distribution that supports Oracle DB 12.2.0.1 to create a rack configuration as if a single cluster were to be installed with all DB servers and cells from both sites. This file will be used to partially configure the Exadata components in a later step.

Install and configure Exadata racks in desired sites

Perform a typical installation of Exadata racks on each site.

Install additional IB cables between each of the Exadata racks to the central spine switch, if one is used (various cabling schemes are possible, see the documentation. The central spine switch must be in a physical location that has independent power, HVAC, etc from the *quorum* site and the other cluster sites.

Use OEDA for preliminary configuration

Use the already created OEDA XML file to perform a basic configuration of the DB servers and cells.

NOTE: Ensure the DB and Grid Homes match what you will create later when performing the manual installation of Grid Infrastructure (GI) and database.

Using OEDA *install.sh,* do the steps **before** attempting the "Install Cluster Software" step. For example, do the following steps (the exact numbers vary depending on OEDA version)::

```
# ./install.sh -cf ./exa01.xml  -l
 Initializing
 -----
1. Validate Configuration File
2. Setup Required Files
3. Create Users
4. Setup Cell Connectivity
5. Verify Infiniband
6. Calibrate Cells
7. Create Cell Disks
8. Create Grid Disks
9. Configure Alerting
```

Create group files and equivalency for DCLI commands

Create group files for DB servers and cells in the first DB node for root; in this example :

dbs_group : exa01adm01, exa01adm02, exa01adm03, exa01adm04

cell_site1 : exa01celadm01, exa01celadm02 exa01celadm03

cell_site2 : exa01celadm04, exa01celadm05 exa01celadm06

all_group: all DB servers and cells from both sites

Ensure that DCLI can be used, create equivalency as needed:

```
# dcli -g ~/all_group -l root date
```

In the above prompts for passwords, then set up equivalency using –k option, for example:

```
# dcli -g ~/all_group -l root -k (answer logon prompts)
```

Copy the dbs_group to the Oracle user's home (and grid user's home if applicable) and ensure dcli can be used:

```
$ dcli -g ~/dbs_group -l oracle date
```

Use the "-k" option as shown in the above command if prompted for a password.

Manually configure NFS server and client

Take note of the oracle user and oinstall userID and groupIDs (typically 1000 and 1001, respectively) after running the OEDA *install.sh* script above. The voting file locations will be owned by this user with the same UID and GIDs across the DB servers and the quorum site. This setup has been tested with OL6 as installed on Exadata compute nodes with the 12.2.1.1.0 image.

On the first DB node, run the following:

```
[root@exa01adm01 u01]# su - oracle
[oracle@exa01adm01 ~]$ id
uid=1001(oracle) gid=1001(oinstall) groups=1001(oinstall),1002(dba)
```

Notice the primary group, GID 1001, is the OINSTALL group, and DBA group has GID 1002. These UID and GID values will be used in this example.

If using role separation then do the following instead (assuming Grid Infrastructure owner is 'grid'):

```
# su - grid
$ id
uid=1000(grid) gid=1001(oinstall) groups=1001(oinstall),1004(asmdba),1005(asmoper),1006(asmadmin)
```

Configure NFS server on the quorum site

The NFS server ideally uses two 10g Ethernet interfaces bonded together.

Perform the following steps on the quorum site as root.

Create the *dba* group with the GID as found in the previous step. The UID and GID must match the ones used on the cluster nodes.

```
# /usr/sbin/groupadd -g 1002 dba
```

Create an oracle user with the UID and GID as found in the previous step

```
# /usr/sbin/useradd -u 1001 -g dba -G dba -d /home/oracle -s /bin/bash oracle
```

If using role separation then do the following instead (assuming Grid Infrastructure owner is 'grid'):

```
# /usr/sbin/groupadd -g 1006 asmadmin
# /usr/sbin/useradd -u 1000 -g asmadmin -G asmadmin -d /home/grid -s /bin/bash grid
```

Make the following changes:

```
# chmod 644 /etc/hosts.allow
# chmod 644 /etc/hosts.deny
```

Comment out "ALL:ALL" in /etc/hosts.deny

NOTE: The above changes are for testing only; production systems should adhere to strict NFS security practices.

Create the voting disk locations

```
# mkdir /u01/quorumdisks
# chown oracle:dba /u01/quorumdisks
```

Or, for role-separated environment:

```
# chown grid:asmadmin /u01/quorumdisks
```

Add the following to the */etc/exports* file:

```
/u01/quorumdisks *(rw,sync)
```

Start the NFS and related daemons:

```
# chkconfig rpcbind on
# chkconfig nfs on
# chkconfig nfslock on
# service rpcbind start
# service nfs start
# service nfslock start
```

Ensure the voting file location is being exported (`qnode01` is the quorum node)

```
# showmount -e qnode01
Export list for qnode01:
/u01/quorumdisks *
```

Configure NFS clients on the DB servers in cluster sites

The choice of network interface to access the NFS server, in order of most preferred to least preferred are:

1. Use two spare 10g Ethernet interfaces bonded together
2. Use the existing bonded client interface, bondeth0
3. Use a single spare 10g Ethernet interface

4. Use the admin network interface

The following configuration steps should be performed on **every** DB node on **all** sites. This can be done with DCLI if you have configured it in a prior step. The group ownership should correspond to the *DBA* group name when installing as non-role separated or the *OSASM* group name that will be used when installing as role-separated (see the OEDA configuration for *ASM Home Admin Group* for the value to use).

Create the voting disk locations as root:

```
# dcli -g ~/dbs_group -l root "mkdir /u01/quorumdisks"
# dcli -g ~/dbs_group -l root "chown oracle:dba /u01/quorumdisks"
```

If using role separation then do the following instead (assuming Grid Infrastructure owner is 'grid' and ASM Home Admin Group is asmadmin):

```
# dcli -g ~/dbs_group -l root "mkdir /u01/quorumdisks"
# dcli -g ~/dbs_group -l root "chown grid:asmadmin /u01/quorumdisks"
```

Add the following to the /etc/fstab file on all DB servers (one continuous line):

```
qnode01:/u01/quorumdisks /u01/quorumdisks nfs rw,bg,soft,
timeo=30,nointr,tcp,vers=3,rsize=32768,wsize=32768,actimeo=0
```

NOTE: Ensure the node name (here, qnode01) is the host name accessed via the client network from the DB servers).

Start the NFS and related daemons:

```
# dcli -g ~/dbs_group -l root "chkconfig rpcbind on"
# dcli -g ~/dbs_group -l root "chkconfig nfslock on"
# dcli -g ~/dbs_group -l root "chkconfig nfs on"

# dcli -g ~/dbs_group -l root "service rpcbind start"
# dcli -g ~/dbs_group -l root "service nfslock start"
# dcli -g ~/dbs_group -l root "service nfs start"
```

Mount the NFS filesystem on all nodes

```
# dcli -g ~/dbs_group -l root "mount /u01/quorumdisks"
```

Sanity check the voting file mount points to ensure they can be used. For example:

```
# df
Filesystem           1K-blocks      Used Available Use% Mounted on
…
```

```
qnode01-:/u01/quorumdisks   103081248  34554688  63267296  36% /u01/quorumdisks

# su – oracle  (or grid if using role separation)
$ cd /u01/quorumdisks
$ touch deleteme
$ dcli –g ~/dbs_group –l oracle "ls -l /u01/quorumdisks/deleteme"
< You should see deleteme from all DB servers >

$ rm deleteme
```

Create the voting files in the new voting file location:

1. You must create as many empty voting files as there are griddisks in each failure group. This will equal the number of disks found on cell for the DATAC1 griddisks. For HC cells of an eighth rack it will be 6 disks, for quarter racks or greater it will be 12 disks; for EF cells this will be 8 disks.

2. Adjust the placeholder in the command below, <NUM of DISKS> to the value determined above. Also, adjust the quorum disk name to be in this pattern: <rack name>_<diskgroup name>_qdisk<DISK NUM>. For example: exa01_DATAC1_qdisk.

   Then, execute the command to create the quorum files (run as *oracle* or as *grid* for role separated environments):

   ```
   $ cd /u01/quorumdisks
   ```

   ```
   $ seq <NUM of DISKS> | awk '{print "dd if=/dev/zero of=<rack name>_<diskgroup name>_qdisk " $1 "
   bs=134217728 count=1 "}' | sh
   ```

   For example: for HC disks (non eighth rack), <NUM of DISKS> is replaced by twelve (12):

   ```
   $ seq 12 | awk '{print "dd if=/dev/zero of=exa01_DATAC1_qdisk" $1 "  bs=134217728 count=1 "}' | sh
   ```

3. You should see the expected number of voting files, for example:

   ```
   $ ls -l
   -rw-rw---- 1 oracle dba 134217728 Feb  6 15:01 exa01h2_DATAC1_qdisk1
   -rw-rw---- 1 oracle dba 134217728 Feb  6 15:01 exa01h2_DATAC1_qdisk2
   …
   -rw-rw---- 1 oracle dba 134217728 Feb  6 15:01 exa01h2_DATAC1_qdisk12
   ```

4. Change the permissions of the voting files to 660
   ```
   $ chmod 660 /u01/quorumdisks/*
   ```

5. Change the ownership to oracle:dba (or *grid:asmadmin* for role separated environments; the group is the one set for OSASM group*)* if not already set:

```
$ chown oracle:dba /u01/quorumdisks/*
```

Manually Configure Exadata Cells

Ensure cells do not have a site_name and site_id already set (repeat for both sites):

```
# dcli -g ~/cell_site1 -l root " cellcli -e list cell attributes name,sitename,siteid"
exa01celadm01: exa01celadm01
exa01celadm02: exa01celadm02
exa01celadm03: exa01celadm03
```

If you need to clear the sitename and siteid:

```
# dcli -g ~/cell_site1 -l root 'cellcli -e alter cell sitename=\"\" '
# dcli -g ~/cell_site1 -l root 'cellcli -e alter cell siteid=\"\"'
```

Give each site their own sitename (siteid will be automatically assigned when the initial sitename is given):

```
# dcli -g ~/cells_site1 -l root 'cellcli -e "alter cell sitename=SITE1"  '
# dcli -g ~/cells_site1 -l root 'cellcli -e "list cell attributes name,sitename,siteid" '
exa01celadm01: exa01celadm01   SITE1   0ef6ffc2-72d5-3885-b713-304e9605b1ce
exa01celadm02: exa01celadm02   SITE1   0ef6ffc2-72d5-3885-b713-304e9605b1ce
exa01celadm03: exa01celadm03   SITE1   0ef6ffc2-72d5-3885-b713-304e9605b1ce

# dcli -g ~/cells_site2 -l root 'cellcli -e "alter cell sitename=SITE2" '
# dcli -g ~/cells_site2 -l root 'cellcli -e "list cell attributes name,sitename,siteid" '
exa01celadm04: exa01celadm04   SITE2   356c390e-b00a-3ecd-a278-08b5aa464c5e
exa01celadm05: exa01celadm05   SITE2   356c390e-b00a-3ecd-a278-08b5aa464c5e
exa01celadm06: exa01celadm06   SITE2   356c390e-b00a-3ecd-a278-08b5aa464c5e
```

NOTE: it is not necessary to pay attention to the *siteid* unless you are troubleshooting an issue. The *siteid* will not change if the *sitename* changes. The important thing is that *siteid* MUST be the same across all cells in a site and the *siteid's* of one site must be different from the other sites.

Manually install Grid Infrastructure

OEDA does not support the installation of extended clusters. Therefore, you must manually install the Grid Infrastructure using the `GridSetup.sh` installation script that comes with the GI software, as shown in the following steps:

1. Ensure VNC is configured and working from the first DB node in the rack that is part of the cluster
2. Obtain the grid_home.zip and database.zip files for 12.2 (placed in this example in $WORK)
3. Create the following grid home directory and unzip the grid_home.zip file:

```
# dcli -g ~/dbs_group -l root mkdir -p /u01/app/12.2.0.1/grid
# dcli -g ~/dbs_group -l root chown oracle:oinstall /u01/app/12.2.0.1/grid
```

If using role separation then do the following instead (assuming Grid Infrastructure owner is 'grid'):

```
# dcli -g ~/dbs_group -l root "chown grid:oinstall" /u01/app/12.2.0.1/grid
```

```
$ unzip -q -d $WORK/grid_home*.zip /u01/app/12.2.0.1/grid
```

4. Start *gridSetup.sh* to install the cluster (use the *oracle user* or the *grid* user if using role separation):

```
$ unset ORACLE_HOME ORACLE_BASE ORACLE_SID
$ export DISPLAY=<your_xserver>:0
$ cd /u01/app/12.2.0.1/grid
$ /u01/app/12.2.0.1/grid/gridSetup.sh
```

NOTE: The major decision points of the installation will be mentioned in the remaining steps. See appendix for details of the installation flow.

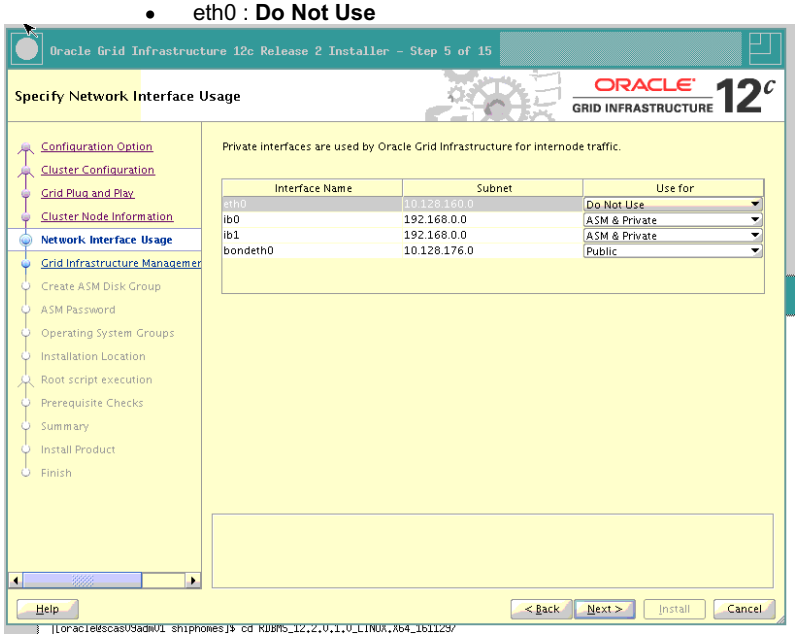
5. Select the option for installing extended cluster. Enter the three site names when asked for them. E.g:

   SITE1, SITE2, SITEQ

6. In the Cluster Node Information dialog, it asks for *Public Hostname.* Supply the administrative hostnames, not the client hostnames. The *Virtual Hostname* is looking for the VIP hostname that resembles, *exa01client01-vip*

7. In the page of selecting the nodes, select Public, Private & ASM interface as in normal clusters:

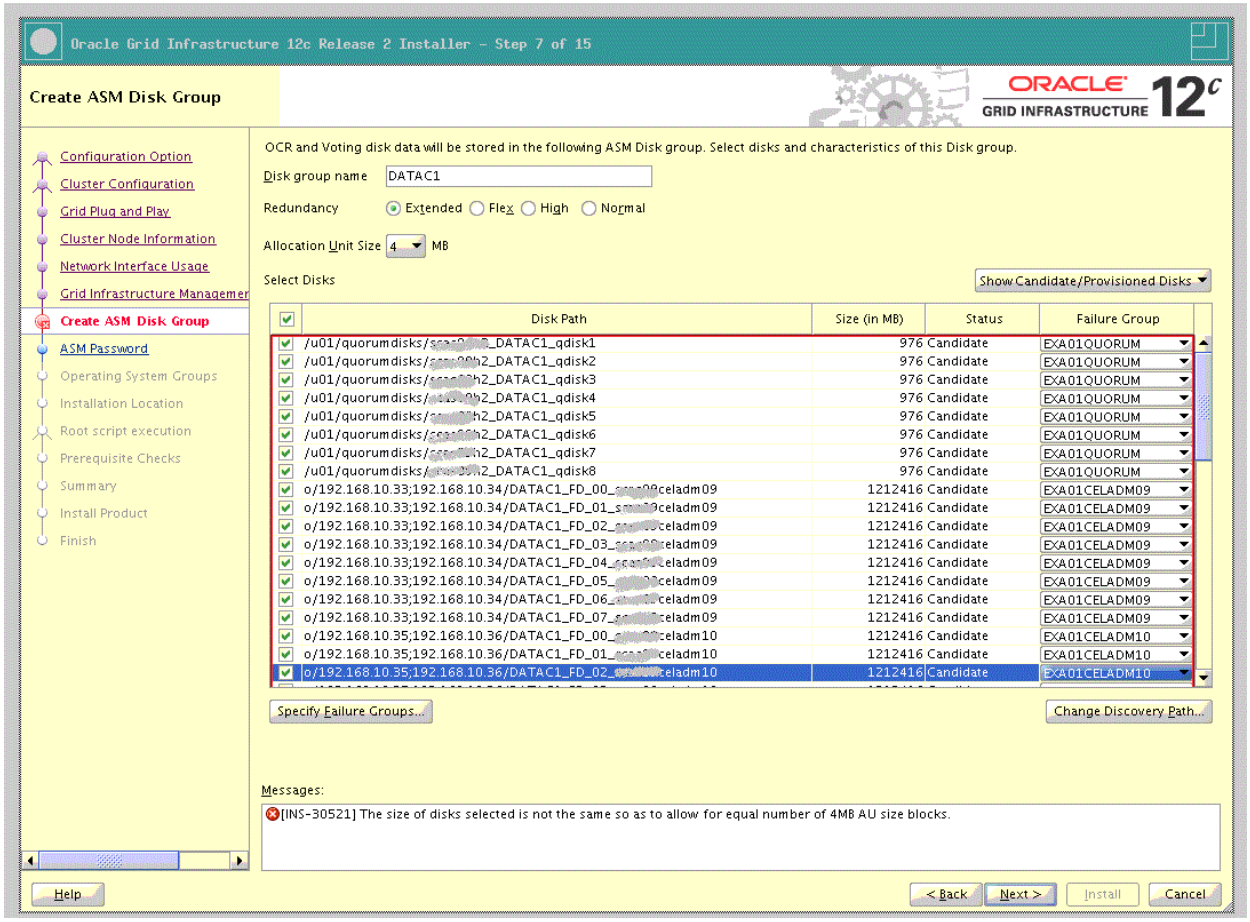
   - bondeth0 (not the admin network): **Public**
   - IB network (ib0 and ib1**): ASM & Private**
   - eth0 : **Do Not Use**

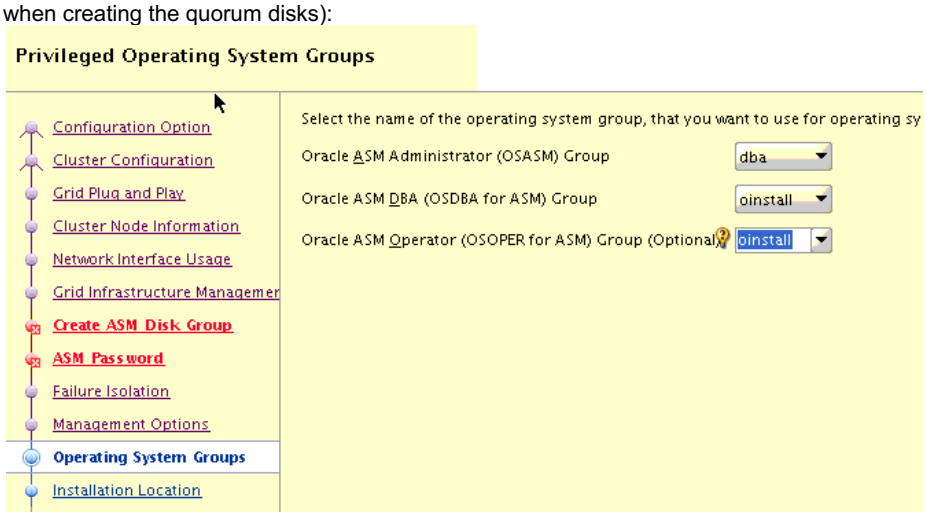


8. Assign proper site names to the db nodes as decided.

9. At the “Create ASM Disk Group” page, do the following:
   - Select extended redundancy for the voting diskgroup
   - Enter the failure group and site information for all the failure groups, it will appear after clicking the button on the lower left, “Specify Failure Groups”. Give names consistent with Exadata, like EXA01CELADM01, EXA01CELADM02, etc. Assign a site name for each failure group. For the quorum failure group, you can name it like EXA01QUORUM or similar.
   - After setting up the failure groups, you return to the main ASM page. Click on the “Change Discovery Path” button on the lower right and set a disk discovery path that will see the griddisks as well as the voting files (its best to remove any existing string that doesn’t apply and show the smallest number of griddisks available), for example:

     ```
     o/DATAC1*/*,/u01/quorumdisks/nfs*
     ```
   - Click on the “Show Candidate/Provisioned Disks” button on the top right to see the disks available. You should see all the voting disks created in NFS and the griddisks created on the cells.
   - Configure the quorum failure group by selecting all of the NFS devices you see (should be equal to the number of disks per cell) and designate them with the failure group name you defined for the quorum site, EXA01QUORUM (or whatever it was set to).
   - Configure the griddisks individually so that they are assigned to the proper failure groups
   - See example image:



- Note, when clicking “Next”, you may see an error like this, "DIsk must be same size..." but it allows continuation. This is a bug but may be fixed already…if not, it should be filed.
- Do NOT use IPMI
- For the OS groups, you can use similar settings as this (ensure the OSASM value is the same as the one set earlier when creating the quorum disks):
- 

On the “Specify Installation Location” page, you may see an error that /u01/app/oracle is not empty on some nodes. It may contain residual OEDA directories placed there during the OEDA installation phase. You can clear these out by running this from the first DB node (assumes you’ve created equivalency with the oracle user):

```
dcli -g ~/dbs_group -l root "rm -rf /u01/app/oracle"
```

or, for role-separated environments:

```
dcli -g ~/dbs_group -l root "rm -rf /u01/app/grid"
```

In addition, run the following commands needed for a role-separated installation:

```
# dcli -g dbs_group -l root 'mkdir -p /u01/app/grid'
# dcli -g dbs_group -l root 'chown grid:oinstall /u01/app/grid'
# dcli -g ~/dbs_group -l root 'chown -R grid:oinstall /u01'
# dcli -g ~/dbs_group -l root 'chmod -R 775 /u01/'
```

- The subsequent pages are similar to non-Extended Cluster installations (see Appendix for details)
- Make sure to save a copy of the response file
- “Perform Prerequisite Checks” may catch some problems…fix those issues and retry; do not proceed unless the issues are addressed.
- Save a response file at the “Summary” page. This could save you from re-entering data if you need to re-run this.
- Start the installation
- Run root.sh when prompted.
  NOTE: DO NOT RUN THEM IN PARALLEL – IT CAN CAUSE OUT-OF-ORDER ASM INSTANCE NAME SEQUENCE…e.g., +ASM4 will be on the third node.
- Finish the installation

After it is finished, check that the cluster is running:

```
# crsctl get cluster extended

CRS-6577: The cluster is extended.

# crsctl stat res –t : Should show expected nodeapps running on all nodes

# crsctl query cluster site -all

Site 'site1' identified by GUID '0ef6ffc272d53885b713304e9605b1ce' in state 'ENABLED'
contains nodes 'exa01adm05,exa01adm06' and disks
'DATAC1_FD_00_EXA01CELADM09,DATAC1_FD_00_EXA01CELADM10,DATAC1_FD_00_EXA01CELADM11,DATAC1_FD_
01_EXA01CELADM09,DATAC1_FD_01_EXA01CELADM10,DATAC1_FD_01_EXA01CELADM11,DATAC1_FD_02_EXA01CEL
ADM09,DATAC1_FD_02_EXA01CELADM10,DATAC1_FD_02_EXA01CELADM11,DATAC1_FD_03_EXA01CELADM09,DATAC
1_FD_03_EXA01CELADM10,DATAC1_FD_03_EXA01CELADM11,DATAC1_FD_04_EXA01CELADM09,DATAC1_FD_04_EXA
01CELADM10,DATAC1_FD_04_EXA01CELADM11,DATAC1_FD_05_EXA01CELADM09,DATAC1_FD_05_EXA01CELADM10,
DATAC1_FD_05_EXA01CELADM11,DATAC1_FD_06_EXA01CELADM09,DATAC1_FD_06_EXA01CELADM10,DATAC1_FD_0
6_EXA01CELADM11,DATAC1_FD_07_EXA01CELADM09,DATAC1_FD_07_EXA01CELADM10,DATAC1_FD_07_EXA01CELA
DM11'.

Site 'site2' identified by GUID '356c390eb00a3ecda27808b5aa464c5e' in state 'ENABLED'
contains nodes 'exa01adm08,exa01adm07' and disks
'DATAC1_FD_00_EXA01CELADM12,DATAC1_FD_00_EXA01CELADM13,DATAC1_FD_00_EXA01CELADM14,DATAC1_FD_
01_EXA01CELADM12,DATAC1_FD_01_EXA01CELADM13,DATAC1_FD_01_EXA01CELADM14,DATAC1_FD_02_EXA01CEL
ADM12,DATAC1_FD_02_EXA01CELADM13,DATAC1_FD_02_EXA01CELADM14,DATAC1_FD_03_EXA01CELADM12,DATAC
1_FD_03_EXA01CELADM13,DATAC1_FD_03_EXA01CELADM14,DATAC1_FD_04_EXA01CELADM12,DATAC1_FD_04_EXA
01CELADM13,DATAC1_FD_04_EXA01CELADM14,DATAC1_FD_05_EXA01CELADM12,DATAC1_FD_05_EXA01CELADM13,
DATAC1_FD_05_EXA01CELADM14,DATAC1_FD_06_EXA01CELADM12,DATAC1_FD_06_EXA01CELADM13,DATAC1_FD_0
6_EXA01CELADM14,DATAC1_FD_07_EXA01CELADM12,DATAC1_FD_07_EXA01CELADM13,DATAC1_FD_07_EXA01CELA
DM14'.
```

Ensure ASM is using Best Practice Settings

Change the Flex ASM configuration to run ASM on all nodes

```
$ srvctl modify asm -count ALL

$ srvctl status  asm -detail

ASM is running on exa01adm06,exa01adm05,exa01adm08,exa01adm07
ASM is enabled.
ASM instance +ASM1 is running on node exa01adm05
Number of connected clients: 2
Client names: -MGMTDB:_mgmtdb:exa01H2-clu exa01adm05.us.oracle.com:_OCR:exa01H2-clu
ASM instance +ASM2 is running on node exa01adm06
Number of connected clients: 1
Client names: exa01adm06.us.oracle.com:_OCR:exa01H2-clu
ASM instance +ASM4 is running on node exa01adm07
Number of connected clients: 1
Client names: exa01adm07.us.oracle.com:_OCR:exa01H2-clu
```

```
ASM instance +ASM3 is running on node exa01adm08
Number of connected clients: 1
Client names: exa01adm08.us.oracle.com:_OCR:exa01H2-clu
```

Change ASM Initialization Parameters

Change the exact values of the IB IP addresses in the commands below to reflect the system being configured and then run the following in SQLPlus while connected to ASM as SYSASM.

It is strongly recommended to create a backup pfile of the spfile before making these changes:

```
create pfile='/home/oracle/asm.pfile' from spfile;
```

NOTE: Check the correct value for cluster_interconnects by looking at ib0 and ib1 on each node:

```
$ dcli -g ~/dbs_group -l root " ip addr show ib0 | grep 'inet ' ;  ip addr show ib1 | grep 'inet '"
exa01adm05: inet 192.168.10.9/20 brd 192.168.15.255 scope global ib0
exa01adm05: inet 192.168.10.10/20 brd 192.168.15.255 scope global ib1
exa01adm06: inet 192.168.10.11/20 brd 192.168.15.255 scope global ib0
exa01adm06: inet 192.168.10.12/20 brd 192.168.15.255 scope global ib1
exa01adm07: inet 192.168.10.13/20 brd 192.168.15.255 scope global ib0
exa01adm07: inet 192.168.10.14/20 brd 192.168.15.255 scope global ib1
exa01adm08: inet 192.168.10.15/20 brd 192.168.15.255 scope global ib0
exa01adm08: inet 192.168.10.16/20 brd 192.168.15.255 scope global ib1
```

Create a file with ALTER SYSTEM commands similar to the ones below to change the parameters. Use SQL*Plus to run this file on one instance.

NOTE: Double check the interconnect IP addresses to ensure they are correct according to the above ip command output

```
ALTER SYSTEM SET cluster_interconnects='192.168.10.9:192.168.10.10' SCOPE=SPFILE SID='+ASM1';
ALTER SYSTEM SET cluster_interconnects='192.168.10.11:192.168.10.12' SCOPE=SPFILE SID='+ASM2';
ALTER SYSTEM SET cluster_interconnects='192.168.10.13:192.168.10.14' SCOPE=SPFILE SID='+ASM3';
ALTER SYSTEM SET cluster_interconnects='192.168.10.15:192.168.10.16' SCOPE=SPFILE SID='+ASM4';
ALTER SYSTEM SET sga_target=3071M SCOPE=SPFILE SID='*';
ALTER SYSTEM SET pga_aggregate_target=400M SCOPE=SPFILE SID='*';
ALTER SYSTEM SET memory_target=0 SCOPE=SPFILE SID='*';
ALTER SYSTEM SET memory_max_target=0 SCOPE=SPFILE SID='*';
ALTER SYSTEM SET processes=1040 SCOPE=SPFILE SID='*';
ALTER SYSTEM SET audit_syslog_level='local0.info' SCOPE=SPFILE SID='*';
ALTER SYSTEM SET audit_sys_operations=TRUE SCOPE=SPFILE SID='*';
ALTER SYSTEM SET asm_power_limit=4 SCOPE=SPFILE SID='*';
```

Restart the cluster on all nodes to make the above settings active.

Verify that voting files include at least one disk from the quorum disk site (NFS):

```
# crsctl query css votedisk
##  STATE    File Universal Id               File Name Disk group
```

```
--  -----    -----------------                --------- ---------
 1. ONLINE   e49179a7c6674f37bfa7008bdaf76938 (o/192.168.10.43;192.168.10.44/DATAC1_FD_06_exa01celadm14) [DATAC1]
 2. ONLINE   503690b82bbf4f14bf5928021c7e86bb (o/192.168.10.39;192.168.10.40/DATAC1_FD_05_exa01celadm12) [DATAC1]
 3. ONLINE   e407d10c51324f60bf9ab72a2e7181cd (o/192.168.10.41;192.168.10.42/DATAC1_FD_06_exa01celadm13) [DATAC1]
 4. ONLINE   a911a4843c584f15bf06456abbc4dafc (o/192.168.10.33;192.168.10.34/DATAC1_FD_06_exa01celadm09) [DATAC1]
 5. ONLINE   f88574e02d434fc2bfa0cbeaf9ee04e8 (o/192.168.10.37;192.168.10.38/DATAC1_FD_05_exa01celadm11) [DATAC1]
 6. ONLINE   1ca077893c014f33bfc4bcaa9442caf2 (o/192.168.10.35;192.168.10.36/DATAC1_FD_06_exa01celadm10) [DATAC1]
 7. ONLINE   40c4f25054604fa4bf2d38d382cdc3c4 (/u01/quorumdisks/exa01h2_DATAC1_qdisk3) [DATAC1]
```

Manually Create a RECO Diskgroup

At this point, we only have one diskgroup, DATAC1. We would like to create an additional diskgroup for the Fast Recovery Area, FRA, commonly called the RECOC1 diskgroup. We will need to create RECOC1 as an EXTENDED REDUNDANCY diskgroup – a typical NORMAL or HIGH redundancy diskgroup will not be able to survive a site failure.

Please do the following:

Create additional quorum disks in the NFS filesystem. Although these will not contain voting files, they are needed by ASM.

1. Create a quorum disk for the new RECO extended redundancy diskgroup (only one quorum disk is needed)

   For example,

   ```
   dd if=/dev/zero of=exa01_RECOC1_qdisk1  bs=134217728 count=1 "}' | sh

   $ ls -l
   -rw-r--r-- 1 oracle oinstall 134217728Jan 20 22:16 reco1
   ```

2. Change the permissions of the voting files to 660

   ```
   $ chmod 660 /u01/quorumdisks/*
   ```

3. Change the ownership to oracle:dba if not already set:

   ```
   $ chown oracle:dba /u01/quorumdisks/*
   ```

   Or for role-separated environment:

   ```
   $ chown grid:asmadmin /u01/quorumdisks/*
   ```

4. Alter the diskstring on all ASM instances to see the new quorum files:

   ```
   ALTER SYSTEM SET
   asm_diskstring='o/*/DATAC1*','o/*/RECOC1*','/u01/quorumdisks/exa01_DATAC1_qdisk*','/u01/qu
   orumdisks/exa01_RECOC1_qdisk*' SCOPE=BOTH SID='*';
   ```

5. Take note of the failure group names used in the DATAC1 diskgroup and use the same names in the next command (should be all uppercase):

   ```
   SQL> select distinct failgroup from v$asm_disk where name is not null;
   ```

6. Create the RECOC1 diskgroup with a command similar to the following (this was done on an EF system, hence only 8 disks per failure group and 8 quorum disks):

```
CREATE DISKGROUP RECOC1 EXTENDED REDUNDANCY
 SITE SITE1  FAILGROUP EXA01CELADM01  DISK 'o/192.168.10.17;192.168.10.18/RECOC1_FD_05_EXA01celadm01'
 SITE SITE1  FAILGROUP EXA01CELADM01  DISK 'o/192.168.10.17;192.168.10.18/RECOC1_FD_04_EXA01celadm01'
 SITE SITE1  FAILGROUP EXA01CELADM01  DISK 'o/192.168.10.17;192.168.10.18/RECOC1_FD_01_EXA01celadm01'
 SITE SITE1  FAILGROUP EXA01CELADM01  DISK 'o/192.168.10.17;192.168.10.18/RECOC1_FD_03_EXA01celadm01'
…
 SITE SITE1  FAILGROUP EXA01CELADM03  DISK 'o/192.168.10.21;192.168.10.22/RECOC1_FD_01_EXA01celadm03'
 SITE SITE1  FAILGROUP EXA01CELADM03  DISK 'o/192.168.10.21;192.168.10.22/RECOC1_FD_05_EXA01celadm03'
 SITE SITE1  FAILGROUP EXA01CELADM03  DISK 'o/192.168.10.21;192.168.10.22/RECOC1_FD_03_EXA01celadm03'
 SITE SITE2  FAILGROUP EXA01CELADM04  DISK 'o/192.168.10.23;192.168.10.24/RECOC1_FD_02_EXA01celadm04'
 SITE SITE2  FAILGROUP EXA01CELADM04  DISK 'o/192.168.10.23;192.168.10.24/RECOC1_FD_00_EXA01celadm04'
 SITE SITE2  FAILGROUP EXA01CELADM04  DISK 'o/192.168.10.23;192.168.10.24/RECOC1_FD_05_EXA01celadm04'
 SITE SITE2  FAILGROUP EXA01CELADM04  DISK 'o/192.168.10.23;192.168.10.24/RECOC1_FD_03_EXA01celadm04'
…
 SITE SITE2  FAILGROUP EXA01CELADM06  DISK 'o/192.168.10.27;192.168.10.28/RECOC1_FD_03_EXA01celadm06'
 SITE SITEQ  QUORUM  FAILGROUP EXA01QUORUM  DISK '/u01/quorumdisks/exa01_RECOC1_qdisk1' SIZE 976M
ATTRIBUTE 'compatible.asm'='12.2.0.1','compatible.rdbms'='12.2.0.1'
,'compatible.advm'='12.2.0.1','au_size'='4M'
,'cell.smart_scan_capable'='TRUE', 'content.type'='recovery';
```

The quorum disk is required for extended redundancy disk groups.

See appendix 2 for a SQL script to generate this command.

7. Mount the RECOC1 diskgroup on all ASM instances

```
$ dcli -g ~/dbs_group -l oracle /u01/app/12.2.0/grid/bin/asmcmd mount RECOC1
```

(use *-l grid* for role-separated environments)

NOTE: this will throw an error that RECOC1 is already mounted on the first node

Check that RECO1 is mounted on all nodes:

```
SQL> select inst_id, name,state from gv$asm_diskgroup order by 1;

   INST_ID NAME                           STATE
---------- ------------------------------ -----------
         1 DATAC1                         MOUNTED
         1 RECOC1                         MOUNTED
         2 RECOC1                         MOUNTED
         2 DATAC1                         MOUNTED
         3 DATAC1                         MOUNTED
         3 RECOC1                         MOUNTED
         4 DATAC1                         MOUNTED
         4 RECOC1                         MOUNTED
```

8. Make sure cell cachingpolicy is properly set for all diskgroups
   - All diskgroups used for data should have their griddisk's cachingPolicy set to "default".
   - All diskgroups used as Fast Recovery Area (typically called "RECOC1") should have their griddisk's cachingPolicy set to "none".

   Confirm this by running the following commands:

```
# dcli -g ~/cell_group -l root "cellcli -e list griddisk attributes
name,cachingpolicy,cachedby,asmdiskgroupname where asmdiskgroupname not like \'RECO.*\' "
```

```
exa01celadm01: DATAC1_FD_00_exa01celadm01      default         FD_00_exa01celadm01    DATAC1
exa01celadm01: DATAC1_FD_01_exa01celadm01      default         FD_01_exa01celadm01    DATAC1
exa01celadm01: DATAC1_FD_02_exa01celadm01      default         FD_02_exa01celadm01    DATAC1
exa01celadm01: DATAC1_FD_03_exa01celadm01      default         FD_03_exa01celadm01    DATAC1
…

# dcli -g ~/cell_group -l root "cellcli -e list griddisk attributes
name,cachingpolicy,cachedby,asmdiskgroupname where asmdiskgroupname like \'RECO.*\' "
exa01celadm01: RECOEXT_FD_00_exa01celadm01     none    FD_00_exa01celadm01    RECOEXT
exa01celadm01: RECOEXT_FD_01_exa01celadm01     none    FD_01_exa01celadm01    RECOEXT
exa01celadm01: RECOEXT_FD_02_exa01celadm01     none    FD_02_exa01celadm01    RECOEXT
…
```

If the *cachingPolicy* needs to be changed, then follow the procedure in the Exadata Server Software documentation.

Use OEDA to complete installation

Use OEDA to complete installation after the cluster is installed and the required diskgroups are created and ready to use. The steps after cluster installation are (the exact numbers vary depending on OEDA version) – notice you will run step 2, then skip to steps 12 through 20, except step 14.

```
 2. Setup Required Files
12. Install Database Software
13. Relink Database with RDS
14. Create ASM Diskgroups ← DO NOT RUN THIS STEP! – as this was already done.
15. Create Databases
16. Create Pluggable Databases [optional]
17. Apply Security Fixes
18. Install Exachk
19. Setup ASR Alerting
20. Create Installation Summary
```

The “Create ASM Diskgroups” step was skipped because that was already done.

Note: After executing Step 15 (Create Databases) of OEDA the DB installation will fail with the following error as seen in the OEDA logs:

```
[FATAL] [DBT-06604] The location specified for 'Fast Recovery Area Location' has insufficient free space.
CAUSE: Only (3,980,610MB) free space is available on the location (+RECOC1/cdbm/).
ACTION: Choose a 'Fast Recovery Area Location' that has enough space (minimum of (4,500,000MB)).
```

The template file corresponding to your DB installation may need to be modified due bug 25512153. For a 2 socket installation, such as X6-2, the template file would be /u01/onecommand/linux-x64/config/OLTP_X2_2_122010.dbt. The change pertains to this entry in the template:

Original: <initParam name="db_recovery_file_dest_size" value="4500000" unit="MB"/>

Modify to: <initParam name="db_recovery_file_dest_size" value="3980500" unit="MB"/>

Validate Configuration

Exachk can be run but it will complain about diskgroup configuration, however there are still valuable checks to review.

Manually ensure that ASM instances have the *asm_diskgroups* parameter including DATAC1, RECOC1, and any other diskgroup you created manually across both sites.

Supporting 11g and 12c Databases

All databases that reside on an extended diskgroup for extended clusters must be 12.2 and the diskgroup must have rdbms.compatible=12.2.0.1. To leverage the same hardware for 11.2 and 12.1 databases, you must create a separate ORACLE_HOME installation for the non-12.2 databases and create a separate set of diskgroups for those databases (typically DATA and RECO) with compatible set to the appropriate RDBMS versions being used (if you plan to use both 11.2 and 12.1 databases with the same diskgroups, you must set rdbms.compatible to the lower version, 11.2). The 11.2 and 12.1 instances cannot span multiple sites; they should use diskgroups defined on the cells local to the database instances.

A customer can create separate physical or virtual clusters (by dividing the rack), or leverage the same 12.2 Grid Infrastructure being used for the extended cluster as mentioned above. For example, your 12.2 databases can be configured in an extended cluster with the following configuration:

Shared Model with one Grid Infrastructure

- 12.2 Grid Infrastructure
    - 12.2 Oracle Home
        - 12.2 DB1 , 12.2 CDB1 (100 PDBs), 12.2 CDB2 (15 PDBs)
        - Sharing the same extended redundancy diskgroups, DATA_EXT and RECO_EXT across 2 sites. 3 storage cells per site
    - 12.1.0.2 Oracle Home
        - 12.1 CDB (25 PDBs), 12.1 DB2
            - Sharing DATA_Site1 and RECO_Site1 local to site 1. Sharing the same storage cells as the 12.2 extended redundancy diskgroups but different grid disks
        - 12.1 CDB (20 PDBs), 12.1 DB3
            - Sharing DATA_Site2 and RECO_Site2 local to site 2. Sharing the same storage cells as the 12.2 extended redundancy diskgroups but different grid disks
    - 11.2.0.4 Oracle Home

- 11.2.0.4 DB4
    - Sharing DATA_Site1 and RECO_Site1 local to site 1. Sharing the same storage cells as the 12.2 extended redundancy diskgroups but different grid disks
- 11.2.0.4 DB5
    - Sharing DATA_Site2 and RECO_Site2 local to site 2. Sharing the same storage cells as the 12.2 extended redundancy diskgroups but different grid disks

Isolated Model with multiple RAC clusters and Grid Infrastructures

- Extended RAC Cluster as above
- RAC Cluster on Site 1 (physical or VM cluster using separate storage disk groups and physical node resources from the extended cluster)
- RAC Cluster on Site 2 (physical or VM cluster using separate storage disk groups and physical node resources resources from the extended cluster)

Best practices for installing non-12.2 RDBMS with 12.2 Grid Infrastructure

An extended cluster infrastructure would be used by 12.2 databases. To run 12.1.0.2 and 11.2.0.4 databases on a 12.2 Grid Infrastructure configured for extended clusters, observe the following best practices:

- Create separate high redundancy diskgroups for the non-12.2 databases at each site.

  For example, if you wanted to install 11.2 databases on site 1 and 12.1 databases on site 2, you would create DATA and RECO diskgroups using griddisks on cells from site 1 for the 11.2 databases and another set of DATA/RECO diskgroups using griddisks on cells from site 2 for the 12.1 databases.

  Ensure each diskgroup has its rdbms.compatible attribute set to accommodate the RDBMS version expected.

- Set the ASM_DISKSTRING on each set of ASM instances per site so that the ASM instances look for disks that are expected for diskgroups at that site in addition to the disks used for extended redundancy diskgroups which span all sites. For example, if site 1 will have diskgroup DATA_112 and RECO_112 composed of griddisks on cells in *site 1* then the ASM_DISKSTRING on *site 1* would have an entry like : `'/o/*/DATA_112*', : '/o/*/RECO_112*'`.

  ASM instances on *site 2* would *not* have their ASM_DISKSTRING include: `'/o/*/DATA_112*', '/o/*/RECO_112*'` but *instead* would have their ASM_DISKSTRING include: `'/o/*/DATA_121*', '/o/*/RECO_121*'` (assuming they have 12.1 diskgroups similar to the other site's 11.2 diskgroups)

- Set the ASM_DISKGROUPS parameter on each set of ASM instances per site so that they attempt to mount only the diskgroups expected to be used at that site. For example, if *site 1* uses DATA_112 and RECO_112 (but NOT DATA_121 and RECO_121),then ensure that ASM_DISKGROUPS only includes DATA_112 and RECO_112 (in addition to the extended redundancy diskgroups that span all sites used by 12.2 instances)

- Configure cluster services to *enable* the 11.2 and 12.1 diskgroups only on the respective sites where they will be used. They should be *disabled* on the sites of the cluster where they will not be used so that Clusterware doesn't try to mount a diskgroup that will only mount for another site. For example, if you have DATA and RECO diskgroups to be used by 11.2 databases on site 1, then you would do the following (assume DATA_112 and RECO_112 are the diskgroup names, site 1 has nodes *exa01adm01* and *exa01adm02 ;* site 2 has nodes *exa01adm03* and *exa01adm04*):

```
srvctl disable diskgroup DATA_112 -node exa01adm03,exa01adm04
```

```
srvctl disable diskgroup RECO_112 –node exa01adm03,exa01adm04
```

Likewise, for DATA_121 and RECO_121 used at site 2 only, you would run the following:

```
srvctl disable diskgroup DATA_121 –node exa01adm01,exa01adm02

srvctl disable diskgroup RECO_121 –node exa01adm01,exa01adm02
```

- Install the 11.2 and 12.1 databases on the DB servers corresponding to the sites where you intend to run them and configure them to use the local diskgroups you created as described above.

    1. On 11.2.0.4 installations, you will need to do the following before running *runInstaller*:

        $GI_HOME_12.2/oui/bin/runInstaller -ignoreSysPrereqs -updateNodeList ORACLE_HOME=$GI_HOME_12.2 "CLUSTER_NODES={exa01adm01,exa01adm02}" CRS=true LOCAL_NODE=exa01adm01

    2. Install using the *software only* option and then patch the newly installed homes according to the recommendations in MOS Doc ID 2234289.1. Alternatively, you can create an OEDA configuration file and use OEDA to install the Oracle software. To do this, you need to do the following:

        - Generate a new configuration file (you can import the one originally used to install 12.2) that specifies the 11.2 or 12.1 home and database.
        - Place the OEDA-required files in the WorkDir (see the OEDA README file)
        - Run *install.sh* steps:

            2. Setup Required Files

            11. Install Database Software

            12. Relink Database with RDS

    3. Use DBCA to create the database, then set *cluster_interconnects* and other parameters according to best practice recommendations

    4. Run *exachk* after installation and configuration is complete to validate that the newly created database adheres to Exadata best practices

Cluster Elasticity

Adding Cluster DB servers

Start with the steps in the *Exadata Database Machine Maintenance Guide, 2.16.1 Adding a New Database Server to the Cluster* . In step 2 when performing the tasks in Re-Imaging Oracle Linux Servers, instead of the indicated procedure for adding the node to the cluster in Task 9, follow the procedure in the *Clusterware Administration and Deployment Guide*, *Using Oracle Grid Infrastructure Installer to Add a Node* for adding a node using gridSetup.sh.

After the node has been added to the cluster, perform *Task 10: Clone Oracle Database Homes To Replacement Database Server* in the *Exadata Database Machine Maintenance Guide.*

Finally, download and run the latest exachk to confirm the node and database were added properly.

Removing Cluster DB servers

Follow the procedure in Deleting Oracle RAC from a Cluster1to remove database instances and RAC from the node.

Remove the node from the cluster using steps in the Clusterware Administration and Deployment Guide, Deleting a Cluster Node on Linux and UNIX Systems.

Adding cells

Adding cells in an extended cluster is similar to adding cells in a standard cluster. After performing tasks listed in the Exadata documentation and after cells are ready to use (but before adding the new cells' griddisks to the existing diskgroup), set the sitename and siteid to match cells in existing sites as follows:

Look up existing sitenames and IDs:

```
# dcli -g ~/cells_site1 -l root 'cellcli -e "list cell attributes name,sitename,siteid" '
exa01celadm01: exa01celadm01    SITE1    0ef6ffc2-72d5-3885-b713-304e9605b1ce
exa01celadm02: exa01celadm02    SITE1    0ef6ffc2-72d5-3885-b713-304e9605b1ce
exa01celadm03: exa01celadm03    SITE1    0ef6ffc2-72d5-3885-b713-304e9605b1ce

# dcli -g ~/cells_site2 -l root 'cellcli -e "list cell attributes name,sitename,siteid" '
exa01celadm04: exa01celadm04    SITE2    356c390e-b00a-3ecd-a278-08b5aa464c5e
exa01celadm05: exa01celadm05    SITE2    356c390e-b00a-3ecd-a278-08b5aa464c5e
exa01celadm06: exa01celadm06    SITE2    356c390e-b00a-3ecd-a278-08b5aa464c5e
```

Set new cells in each site to use the existing sitename and siteid (group files were created for the *new* cells as cells_site1_new and cells_site2_new):

```
# dcli -g ~/cells_site1_new  -l root 'cellcli -e alter cell sitename=\"SITE1\" '
exa01celadm07: Cell EXA01celadm07 successfully altered
exa01celadm08: Cell EXA01celadm08 successfully altered
exa01celadm09: Cell EXA01celadm09 successfully altered
```

```
# dcli -g ~/cells_site1_new  -l root 'cellcli -e alter cell siteid=\"0ef6ffc2-72d5-3885-b713-304e9605
exa01celadm07: Cell EXA01celadm07 successfully altered
exa01celadm08: Cell EXA01celadm08 successfully altered
exa01celadm09: Cell EXA01celadm09 successfully altered

# dcli -g ~/cells_site1_new -l root 'cellcli -e "list cell attributes name,sitename,siteid" '
exa01celadm07: exa01celadm07   SITE1   0ef6ffc2-72d5-3885-b713-304e9605b1ce
exa01celadm08: exa01celadm08   SITE1   0ef6ffc2-72d5-3885-b713-304e9605b1ce
exa01celadm09: exa01celadm09   SITE1   0ef6ffc2-72d5-3885-b713-304e9605b1ce
```

Repeat for SITE2 (site 2 will have different values for the siteid, as shown above).

Add new griddisks into existing diskgroup:

```
ALTER DISKGROUP DATAC1 ADD
SITE SITE1  FAILGROUP EXA01CELADM07  DISK 'o/192.168.10.23;192.168.10.24/DATAC1_FD_00_exa01celadm07'
…
SITE SITE1  FAILGROUP EXA01CELADM07  DISK 'o/192.168.10.23;192.168.10.24/DATAC1_FD_07_exa01celadm07'
SITE SITE2  FAILGROUP EXA01CELADM08  DISK 'o/192.168.10.25;192.168.10.26/DATAC1_FD_00_exa01celadm08'
…
SITE SITE2  FAILGROUP EXA01CELADM08  DISK 'o/192.168.10.25;192.168.10.26/DATAC1_FD_07_exa01celadm08'
REBALANCE POWER 64
```

Note: The above commands assume that the griddisk sizes for the new disks are exactly the same size as the existing *ASM* disk size. If the new griddisks are created to match the existing griddisks, but the existing ASM disk sizes are smaller than the existing griddisk sizes (possible if a resize effort started but didn't finish), the command will throw an ORA- 15410 error.

Creating New diskgroups

See "Manually Create a RECO Diskgroup" above for same procedure. To obtain the benefits of an extended cluster, you must create diskgroups as *extended redundancy*. If you create a normal or high redundancy diskgroup, they will not stay mounted if an entire site goes down.

Resizing Diskgroups

Resizing diskgroups is the same procedure as standard diskgroups except that due to bug 24468470, you must work around the error using a syntax that specifies the failure group and disk size, as follows:

```
ALTER DISKGROUP DBFS_DG RESIZE DISKS IN FAILGROUP
EXA01CELADM01 SIZE 34608M
,EXA01CELADM02 SIZE 34608M
,EXA01CELADM03 SIZE 34608M
,EXA01CELADM04 SIZE 34608M
,EXA01CELADM05 SIZE 34608M
,EXA01CELADM06 SIZE 34608M
,EXA01CELADM07 SIZE 34608M
,EXA01CELADM08 SIZE 34608M
REBALANCE POWER 64;
```

Dropping a Storage Cell from the Cluster

Follow the steps in the Exadata Database Machine Maintenance Guide, *3.6.3 Dropping a Storage Cell from an Existing Disk Group or Storage Grid*.

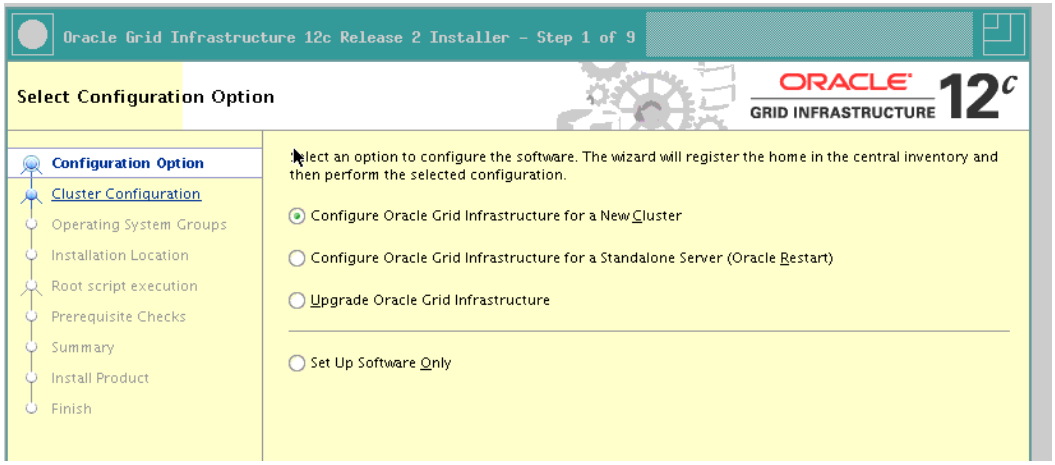
Changing Redundancy Property of an Extended Redundancy Diskgroup

An extended redundancy diskgroup is a type of flex diskgroup. This means it is able to utilize various new features of flex disksgroups such as different redundancy types (MIRROR, 2 copies or HIGH, 3 copies).

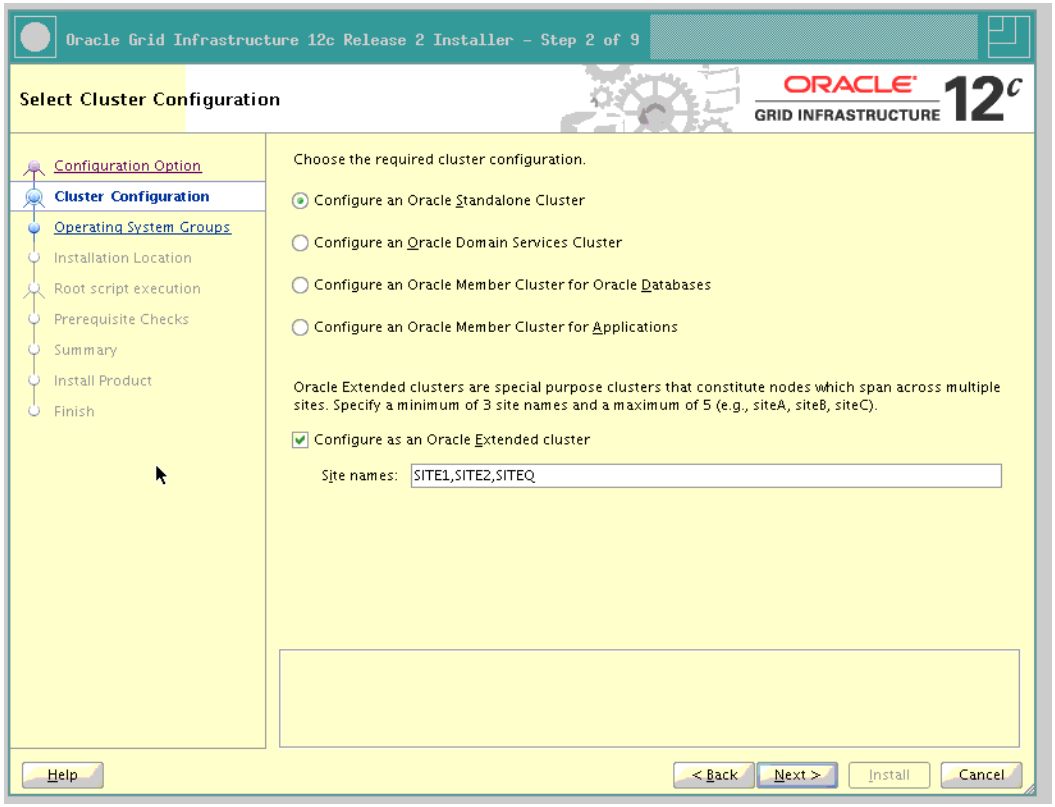
See the ASM documentation for changing redundancy in ASM File Groups.

Appendix: Grid Infrastructure Installation Details

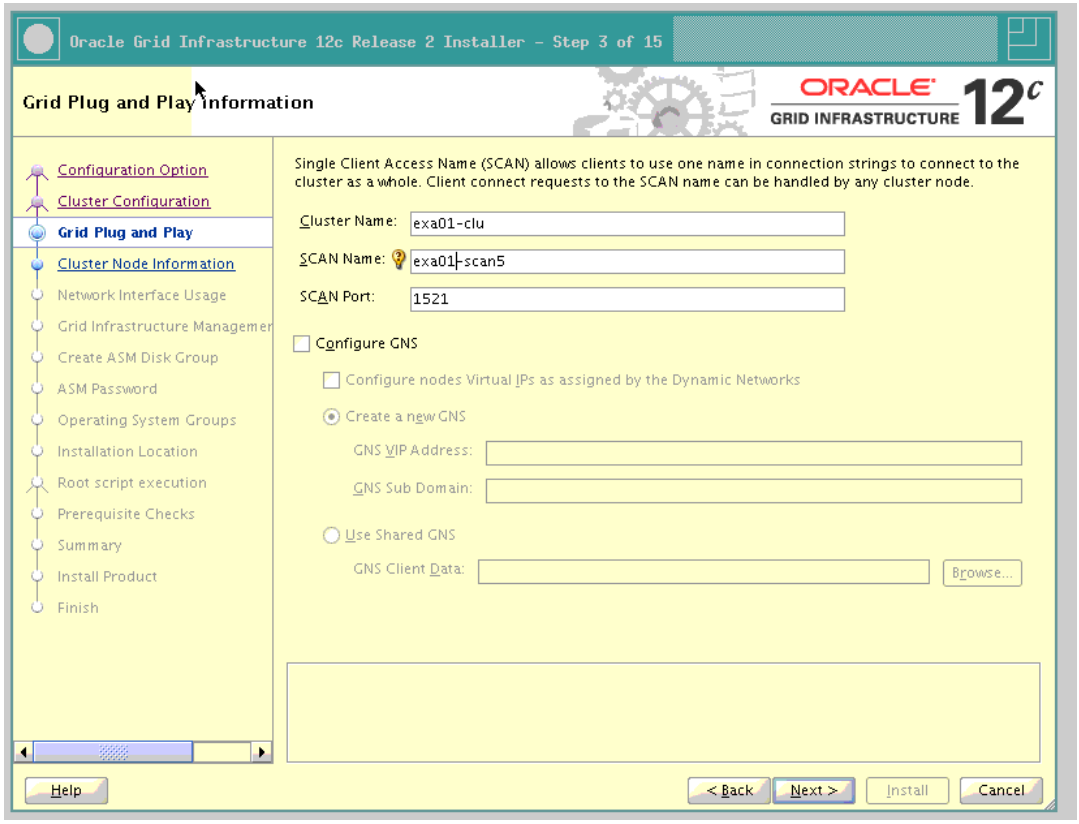
Choose Configuration Option



Choose Extended Clusters and Site Names



Grid Plug N Play

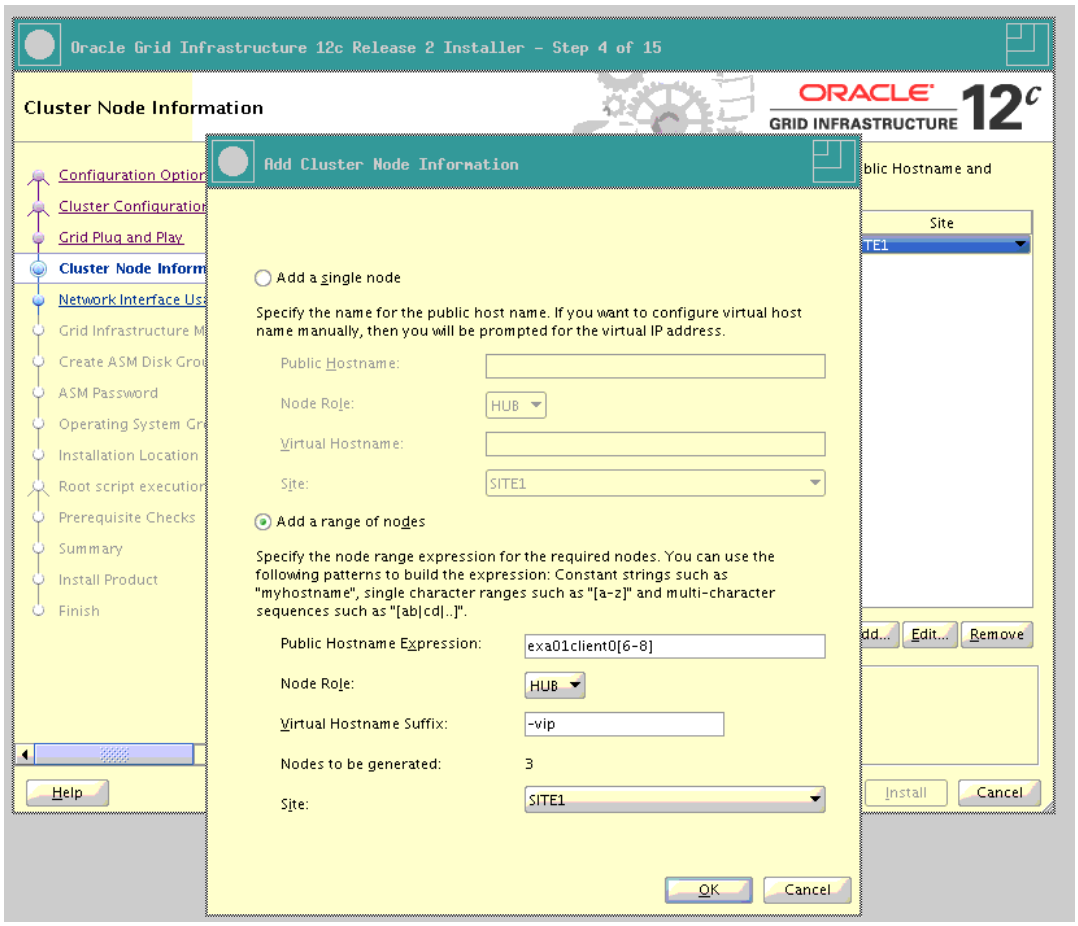


Cluster Node Info

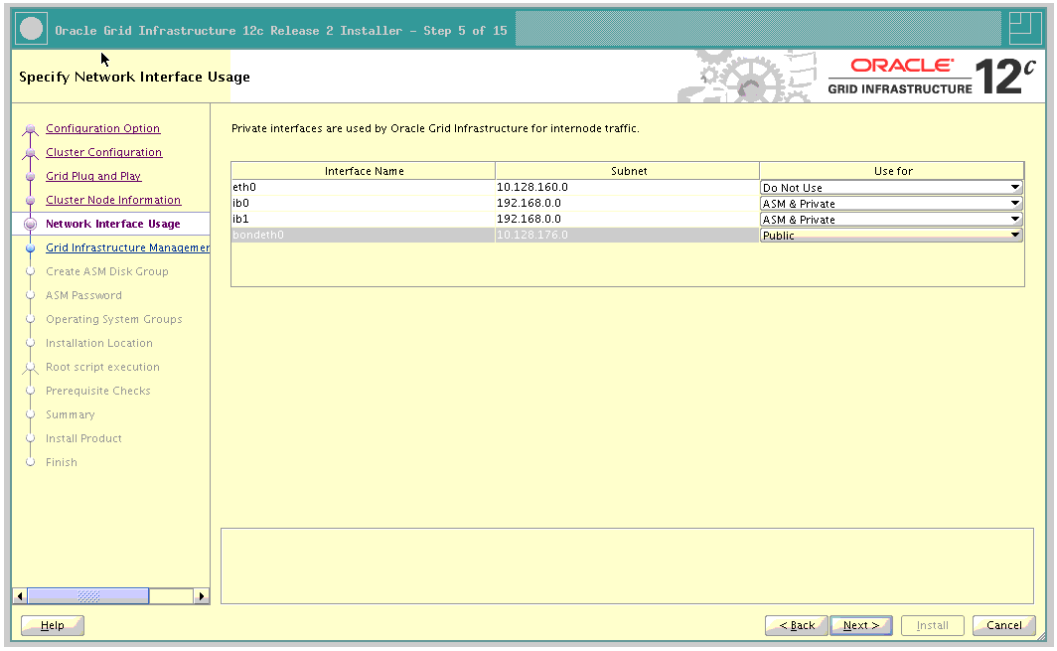
Typically you will need to correct the initially created entry for the first node because the node name should be set to the CLIENT node name, like exa01client01 and not the admin network name.

Then, you will add the remaining nodes using a pattern if possible to make it easy on yourself.

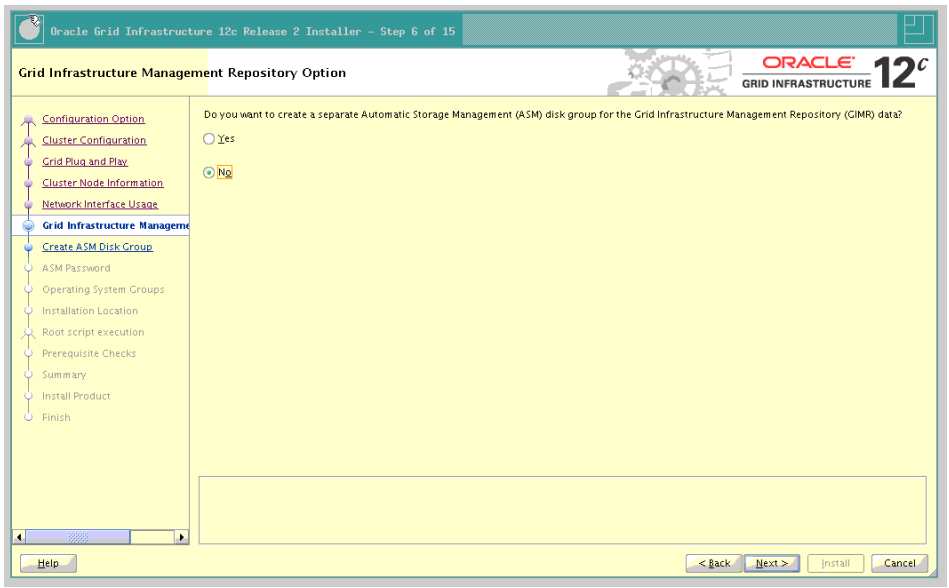
Finally, you will need to re-assign some of the nodes to the correct site.



Network Interface

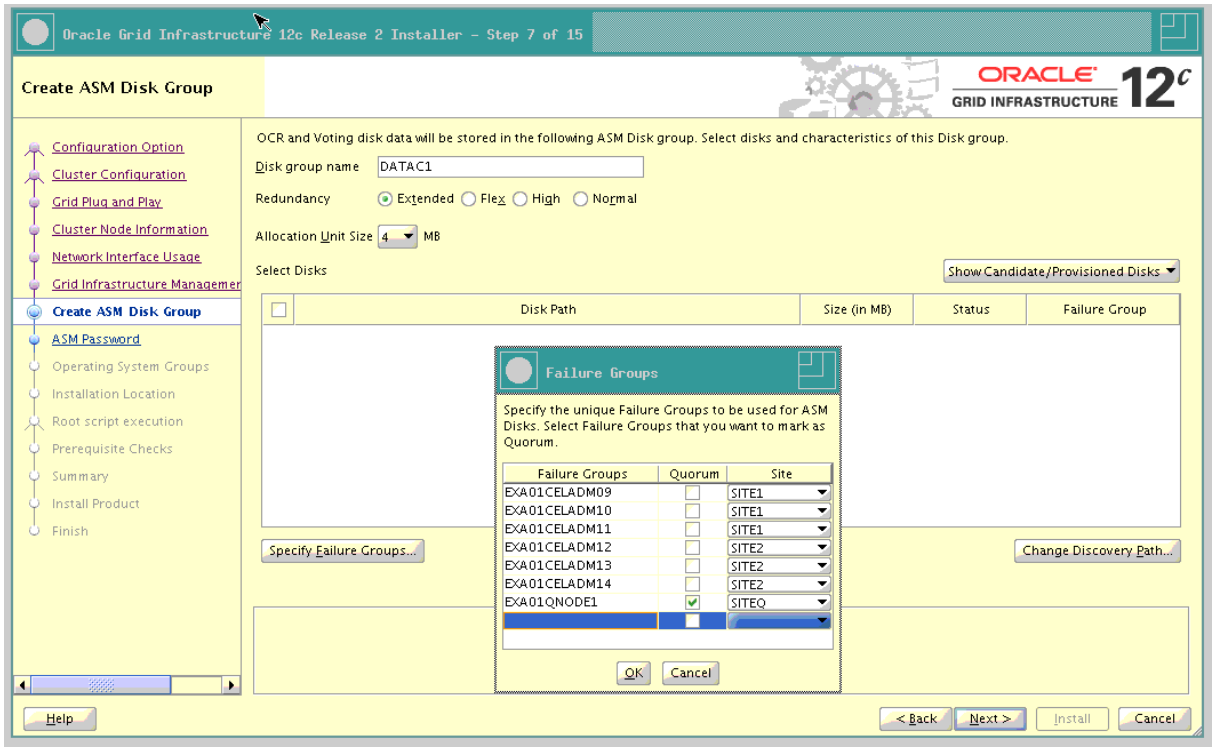


Grid Infrastructure Management Repository



Create ASM Diskgroups

a. Specify the Failure Groups

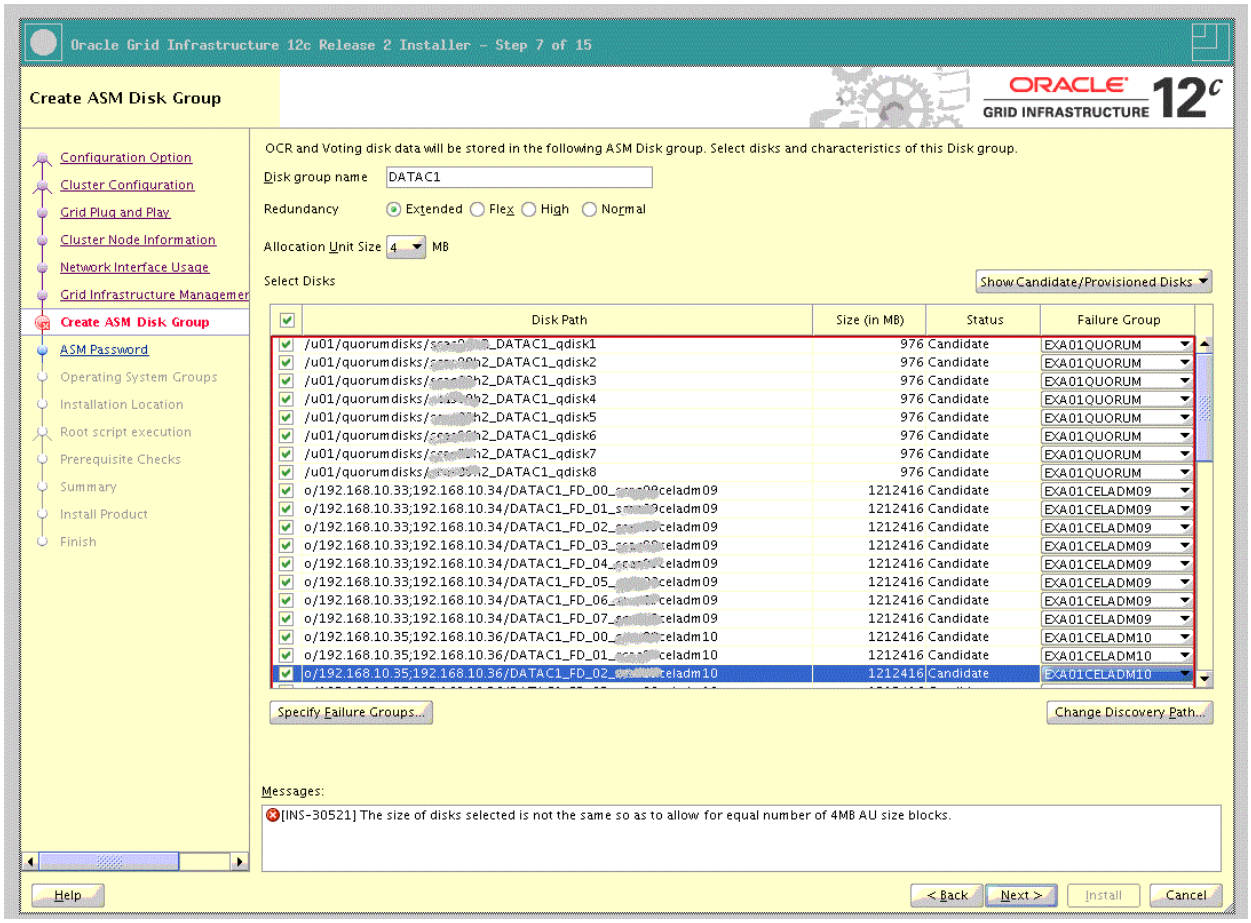


b. Change Discovery Path

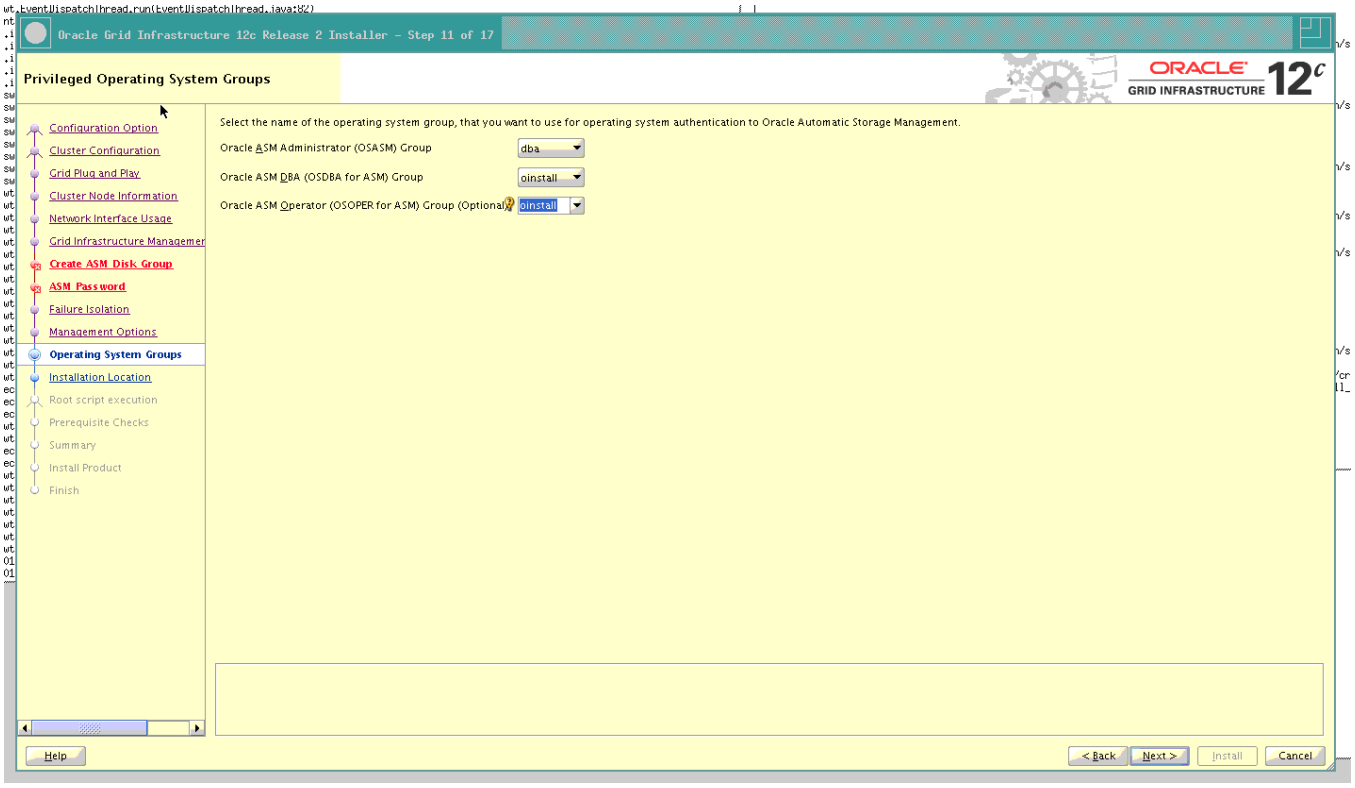
Less is more, only discover the DATAC1 griddisk and the quorum disks

c. Set the Failure Groups to each disk (node names redacted for privacy)

Carefully assign each disk to the appropriate failure group you created in step (a). The error at the bottom of the screen can be ignored and won't impede progress. If there is an error that a mismatch in the number of disks is present, they you have to see which disks have been incorrectly assigned to a failure group.



OS Groups


Oracle Grid Infrastructure 12c Release 2 Installer - Step 11 of 17
Privileged Operating System Groups
ORACLE GRID INFRASTRUCTURE 12c
Configuration Option
Cluster Configuration
Grid Plug and Play
Cluster Node Information
Network Interface Usage
Grid Infrastructure Management
Create ASM Disk Group
ASM Password
Failure Isolation
Management Options
Operating System Groups
Installation Location
Root script execution
Prerequisite Checks
Summary
Install Product
Finish
Select the name of the operating system group, that you want to use for operating system authentication to Oracle Automatic Storage Management.
Oracle ASM Administrator (OSASM) Group
dba
Oracle ASM DBA (OSDBA for ASM) Group
oinstall
Oracle ASM Operator (OSOPER for ASM) Group (Optional)
oinstall
Help
< Back
Next >
Install
Cancel

Appendix 2: Script to Generate Create Diskgroup Script

```
REM
REM Script to generate a create diskgroup command for an extended redundancy RECO diskgroup
REM
REM Run this after you have created the quorum disks and the griddisks for this new diskgroup
REM Replace specific paths and diskgroup name with the actual ones in your environment...look for
REPLACE word
REM Put these commands into a file and run from SQLPlus while logged on to an ASM instance.
REM
REM This script will create a file called create_reco1_diskgroup.sql

spool create_reco1_diskgroup.sql

set heading off
set feedback 0
set verify off
set auto off
set echo off
set lines 400
set pages 0
set trimspool on

select 'CREATE DISKGROUP RECOEXT EXTENDED REDUNDANCY'
    from dual
/
select 'SITE SITE1 FAILGROUP ' || upper(FAILGROUP)
                               || ' DISK '''
                               ||PATH||''''
    from  v$asm_disk
    where path like '%RECOEXT%exa01celadm01'  --REPLACE WITH SITE1 CELL 1
       or path like '%RECOEXT%exa01celadm02'  --REPLACE WITH SITE1 CELL 2
       or path like '%RECOEXT%exa01celadm03'  --REPLACE WITH SITE1 CELL 3
    order by failgroup,path
/
select 'SITE SITE2 FAILGROUP ' || upper(FAILGROUP)
                               || ' DISK '''
                               ||PATH||''''
    from v$asm_disk
    where path like '%RECOEXT%exa01celadm04'  --REPLACE WITH SITE2 CELL 4
       or path like '%RECOEXT%exa01celadm05'  --REPLACE WITH SITE2 CELL 5
       or path like '%RECOEXT%exa01celadm06'  --REPLACE WITH SITE2 CELL 6
    order by failgroup,path
/
select 'SITE SITEQ QUORUM FAILGROUP',(select failgroup
```

```
                                        from v$asm_disk
                                       where path='/u01/quorumdisks/exa01_DATAC1_qdisk1') --REPLACE
PATH WITH PATH OF DATAC1 QUORUM DISK CREATED EARLIER JUST TO GET THE FAILGROUP
                                    ||' DISK '''
                                    ||PATH
                                    ||''''
                                    ||' SIZE 128M'
    from v$asm_disk
    where path like '%RECOEXT_qdisk%'
    order by path
/
select 'ATTRIBUTE'||chr(10)
                  ||'''compatible.asm''=''12.2.0.1'','
                  ||chr(10)
                  ||'''compatible.rdbms''=''12.2.0.1'','
                  ||chr(10)
                  ||'''compatible.advm''=''12.2.0.1'','
                  ||chr(10)
                  ||'''au_size''=''4M'','
                  ||chr(10)
                  ||'''cell.smart_scan_capable''=''TRUE'','
                  ||chr(10)
                  ||'''content.type''=''recovery'';'  --REPLACE APPROPRIATE VALUE FOR CONTENT.TYPE
    from dual
/

spool off
!echo "spool file generated..."
EXIT 0
```

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

ORACLE

CONNECT WITH US

blogs.oracle.com/oracle
facebook.com/oracle
twitter.com/oracle
oracle.com

Integrated Cloud Applications & Platform Services

Installing Oracle Extended Clusters on Exadata
June 2017
Author: Hector Pujol
Contributing Authors: Naren Juneja, Liwen Huang, Lawrence To