

Mise en Œuvre RAC

Un cluster RAC est constitué de :

- **Cluster nodes** – 2 à n nœuds ou serveurs, exécutant chacun un serveur de bases de données Oracle.
- **Network Interconnect** – un réseau privé utilise pour les communications du cluster et du cache fusion. Il est utilisé pour le transfert de blocs de données entre les instances des nœuds.
- **Stockage partagé** – utilisé pour contenir le système de base de données et les fichiers de données. Le stockage partagé est accessible par les nœuds du cluster.
- **Réseau de production (publique)** – utilisé par les clients et les serveurs d'applications pour accéder à la base de données.

Pour construire un RAC :

1. Ajouter un disque (partagé)
2. Partitionner ce disque (minimum 3 diskgroup => 1 pour CRS files : OCR+Voting files ; 1 pour DATA ; 1 pour FRA)
3. Créer les disques ASM (CRS1, DATA1, FRA1) après avoir configuré et initialisé oracleasm.

Le fait d'utiliser un disque partagé sur les différents nœuds évite de partitionner le disque sur chacun des nœuds et aussi évite de créer les disques ASM sur chacun des nœuds.

A quel moment définit-on les IP privées et publiques ?

Lors de l'installation du logiciel clusterware(Oracle Grid Infrastructure).
Le premier diskgroup à créer est celui qui contient les fichiers du cluster (CRS). Il est créé lors de l'installation de Oracle Grid Infrastructure.

1 - Quels sont les prérequis (matériels) pour installer Oracle RAC ?

- Les nœuds doivent être sur le même VLAN ; chaque nœud doit avoir 3 IP : publique, privée et VIP ;
- Le stockage doit être partagé (SAN) entre les nœuds avec de l'ASM ou de l'OCFS
- Un utilisateur oracle pour le compte oracle et un utilisateur grid pour l'installation de Oracle Clusterware
- L'OCR et les voting disks doivent être dans le même diskgroup
- Choix de la redondance (normal pour les données DG_DATA, externally pour la FRA DG_FRA)

En redondance normal il y a 2 failure groups. En redondance externe, il n'y en a pas. En redondance élevée il y en a 3.

Sur les failure groups (groupe d'échecs)... <https://docs.oracle.com/database/121/OSTMG/GUID-378F26F8-00E2-4058-9AE1-D312633D4D81.htm>

Il y a toujours des failure groups même s'il ne sont pas explicitement créés. Des failures groups de capacité inégale, peuvent entraîner des problèmes d'allocation qui empêchent la pleine utilisation de tout le stockage disponible.

Pour un petit nombre de disques (<20), il est généralement préférable d'utiliser la création de failure group par défaut qui place chaque disque dans son propre failure group.

Comment spécifier un failure group ?

```
CREATE DISKGROUP DATA NORMAL REDUNDANCY FAILGROUP DATA_FAILURE_group_1 DISK
'/dev/5855d' , '/dev/6476d' ,
FAILGROUP DATA_FAILURE_group_2 DISK '/dev/5853d' , '/dev/5854d' ;
```

2 - Quels sont les composants critiques dans un RAC

- Les votings disques : Les disques de vote sont des fichiers qui gèrent les informations sur l'appartenance aux nœuds. Ils se trouvent sur le système de cluster partagé ou un fichier de périphérique brut partagé ou Oracle ASM
- L'OLR (oracle local registry) : sous \$GRID_HOME/cdata/*olr
- L'OCR (oracle cluster registry) : L'emplacement de l'OCR est indiqué dans /etc/oracle/ocr.loc.
- Cluster Synchronization Services – CSS : maintient la synchronisation entre l'ASM et les instances de la base de données.

Les process CSS écrivent dans les voting disks dans un block réservé au nœud dont chacun est issu.

Pour reconfigurer CSS suite au changement d'OracleHome de l'ASM :

```
$ORACLE_HOME/bin/localconfig reset
```

- L'interconnect
- Le cache fusion : Dans certains cas, l'instance demande à une autre instance un bloc de données. Elle peut y accéder via le mécanisme d'interconnexion. Le fusion de cache est implémenté par un mécanisme de contrôle appelé Global Cache Service [GCS] . Maintenir l'intégrité des données est une tâche de GCS.

On peut dire que GCS est au cœur du processus du CACHE FUSION. GCS est disponible dans RAC en tant que processus d'arrière-plan LMS.

3 -Quels sont les pré requis de migration d'un ASM d'un SAN vers un autre SAN avec zero downtime?

1. Le nouveau SAN doit être visible : monter les nouveaux disques sur les nœuds RAC
2. Créer les nouveaux disques ASM à partir des disques montés
3. Ajouter les nouveaux disques dans le diskgroup existant
4. Rebalance du storage
5. Supprimer les anciens disques du diskgroup
6. Rebalance du storage

4 - Comment on gère un groupe de disque full rempli à + 90% ?

Il faut ajouter des disques dans le même diskgroup et ensuite faire le rebalancing.

1. Récupérer le N° de LUN à utiliser.
2. Faire découvrir les LUN nouvellement ajoutés aux serveurs.
3. Retrouver le path système correspondant à ces LUN.
4. Identifier le WWID
5. Modification de la configuration multipath
6. Redémarrage du service multipath
service multipathd restart
7. Création des disks
8. Ajout des disk aux diskgroup (à faire sur un seul nœud) :

```
ALTER DISKGROUP DISKGROUPA add failgroup FGA disk  
'ORCL:DISK2' REBALANCE power 3;
```

5 - Différence entre un dataguard physical et logical ?

Pour une physical standby :

- Les schémas d'une base de données Physical standby correspondent exactement à ceux de la base de données source.
- Les archives logs sont directement envoyés sur la base de données de secours qui s'exécute toujours en mode «recover». À l'arrivée, les archives logs sont appliqués directement à la base de données de secours.

Pour une logical standby :

- La base de données de secours logique ne doit pas nécessairement correspondre à la structure de schéma de la base de données source.
- La base de données de secours logique utilise les techniques de LogMiner pour transformer les archives logs en traitements DML natifs (insert, update, delete). Ce DML est transporté et appliqué à la base de données standby.
- La base de données de secours logique peut être ouverte pour des requêtes SQL en lecture seule et toutes les autres tables de secours peuvent être ouvertes pour des mises à jour.
- La base de données de secours logique peut avoir des vues matérialisées et des indexes pour gagner en performance.

6-Comment construire une logical standby ?

A partir d'une base de données physical standby.

- Sur la primary : `exec dbms_logstdby.build`
- Sur la physical : `alter database recover to logical standby testdg;`

Création d'un disque ASM :

```
root # oracleasm createdisk DATA2 /dev/sdc1
```

Ajout d'un disque ASM dans le diskgroup DATA :

```
Alter diskgroup DATA add disk '/dev/oracleasm/disks/DATA2' rebalance power 20;
```

Suppression d'un disque ASM dans le diskgroup DATA :

```
Alter diskgroup DATA drop disk 'DATA_0000' rebalance power 20;
```

Suppression du disque systeme :

```
Root # oracleasm deletedisk DATA1 /dev/sb1
```

Quel est le nombre max que peut accueillir un cluster Oracle ?

En 11G 100 nœuds, en 12c 255 avec OCFS 2.