

Oracle RAC : dépannage du cluster



par Fabien Celaia ([Fadace](#))

Date de publication : 25.11.2006

Dernière mise à jour : 21.3.2008

Procédure manuelle de décatalogage et de recatalogage d'une base dans un OCR/RAC Oracle

I - Historique et glossaire.....	3
II - Problématique.....	4
III - Enjeu et contraintes de production.....	5
III-A - Restauration de l'OCR.....	5
III-B - Backup de l'OCR.....	6
IV - Procédure manuelle.....	7
V - Recréation complète d'un OCR corrompu.....	9

I - Historique et glossaire

*Merci à l'aimable participation de **bouyao** pour ce petit historique*

Historiquement, Oracle fait confiance au vendeur d'OS de fournir la couche clusterware afin d'utiliser Oracle Parallel Server (OPS), et dernièrement RAC. Le DBA devrait demander à l'administrateur OS de configurer l'OS et de créer le cluster en utilisant le logiciel fourni par le vendeur OS.

Depuis la 7.3.3.0.1 sur les plates-formes Windows, ça commence à changer. Par soucis de la nécessité, à cause de la manque d'un logiciel viable de clustering sous windows, Oracle commence à fournir une couche de clustering sur les plates-formes Windows pour utiliser de l'OPS. Par la suite, avec la croissance de la popularité de Linux, Oracle a été obligé de fournir le clusterware sous linux. Avant la 9i, le clusterware a été distribué par les vendeurs OS. Depuis Oracle 9i, le clusterware a été intégré comme une part entière d' Oracle 9i Edition Entreprise.

Avec Oracle 10g, Oracle a introduit le CRS, supporté sur la majorité des OS (Sun, HP, Tru64, AIX, Windows, etc). Quand on utilise Oracle CRS, il est possible d'exécuter le RAC dans un Oracle Edition Standard.

Le CRS est constitué de trois composants majeurs, qui se manifestent comme des deamons (inittab pour Unix et les Services sous Windows). Il existe trois deamons : OCSSD (cluster synchronization services daemon), CRSD qui gere la disponibilité des ressources et EVMD (event logger daemon).

OCSSD et EVMD se lancent sous le compte Oracle et CRSD sous le compte root.

Acronyme	Description
CRS	Cluster Ready Services
RAC	Real Application Cluster
OCR	Oracle Cluster Registry
OCFS	Oracle Cluster File System

II - Problématique

Nous sommes en présence d'un environnement Oracle en RAC, disposé sur 2 noeuds. 3 bases Oracle sont présentes sur ces 2 noeuds. Nous avons donc, en chiffres

- 1 RAC
- 2 noeuds (2 machines) : DVPPORA1 et DVPPORA2
- 3 bases de données Oracle (REP1P, DVP1P et DMP1P)
- 6 instances gérant ces 3 bases (REP1P1 et REP1P2, DVP1P1 et DBVP1P2, DMP1P1 et DMP1P2)
- 3 services TAF (REP1P1, DVP1P1, DMP1P1)
- 2 listeners, un pour chaque machine (LISTENER_DVPPORA1 et LISTENER_DVPPORA2)

Suite à un redémarrage du serveur AIX sur l'un des 2 noeuds du cluster, on constate les statuts suivants :

```
oracle> crs_stat -t
```

Name	Type	Target	State	Host
ora....1P1.srv	application	ONLINE	ONLINE	DVPPORA1
ora....1P2.srv	application	ONLINE	ONLINE	DVPPORA2
ora....B1P1.cs	application	ONLINE	ONLINE	DVPPORA2
ora....P1.inst	application	ONLINE	ONLINE	DVPPORA1
ora....P2.inst	application	ONLINE	ONLINE	DVPPORA2
ora.DVP1P.db	application	ONLINE	ONLINE	DVPPORA2
ora....1P1.srv	application	ONLINE	OFFLINE	
ora....1P2.srv	application	ONLINE	OFFLINE	
ora....P1P1.cs	application	ONLINE	OFFLINE	
ora....P1.inst	application	ONLINE	OFFLINE	
ora....P2.inst	application	ONLINE	OFFLINE	
ora....1P1.srv	application	ONLINE	ONLINE	DVPPORA1
ora....1P2.srv	application	ONLINE	ONLINE	DVPPORA2
ora....P1P1.cs	application	ONLINE	ONLINE	DVPPORA2
ora....P1.inst	application	ONLINE	ONLINE	DVPPORA1
ora....P2.inst	application	ONLINE	ONLINE	DVPPORA2
ora.REP1P.db	application	ONLINE	ONLINE	DVPPORA2
ora....A1.lsnr	application	ONLINE	ONLINE	DVPPORA1
ora....RA1.gsd	application	ONLINE	ONLINE	DVPPORA1
ora....RA1.ons	application	ONLINE	ONLINE	DVPPORA1
ora....RA1.vip	application	ONLINE	ONLINE	DVPPORA1
ora....A2.lsnr	application	ONLINE	ONLINE	DVPPORA2
ora....RA2.gsd	application	ONLINE	ONLINE	DVPPORA2
ora....RA2.ons	application	ONLINE	ONLINE	DVPPORA2
ora....RA2.vip	application	ONLINE	ONLINE	DVPPORA2

Le service DVP1P.db n'apparaît plus dans la liste des services de l'OCR. Il s'agit là d'un problème connu de corruption de l'OCR par le redémarrage d'un des noeuds.

III - Enjeu et contraintes de production

Pour bien faire, il nous faudrait repartir d'une sauvegarde de l'OCR. Ceci nécessiterait malheureusement l'arrêt total des 2 noeuds, et donc une interruption de services des 6 instances (4 saines et 2 mauvaises), ce qui est souvent inacceptable en contrainte de production.

En l'état:

- Il nous est impossible de démarrer la base en standalone sur l'un des 2 noeuds, celle-ci étant toujours référencée dans l'OCR
- Le décataloguage de la base marche, mais le recataloguage est impossible car les 2 instances semblent toujours être cataloguées, bien qu'aucun processus Oracle concernant cette base ne soit présent sur aucun des 2 noeuds (via `ps ax | grep oracle`)
- Le processus d'écoute (commun à toutes les instances d'un même noeud) fonctionne correctement, le tnsping fonctionne.

Juste pour information, la procédure de restauration de l'OCR est décrit ici, mais ce ne sera pas la solution appliquée dans notre cas

III-A - Restauration de l'OCR

 **Interruption de service !!!**

Arrêt du CRS sur tous les noeuds (en tant que root ou via commande sudo)

```
oracle> sudo crsctl stop crs
oracle> sudo /usr/sbin/slibclean
```

Localisation les backups OCR via ocrconfig

```
oracle> ocrconfig -showbackup

DVPORA1      2006/04/27 09:00:35      /oracle/crs10gr2/cdata/crs10gr2
DVPORA1      2006/04/27 05:00:35      /oracle/crs10gr2/cdata/crs10gr2
DVPORA1      2006/04/27 01:00:34      /oracle/crs10gr2/cdata/crs10gr2
DVPORA1      2006/04/25 17:47:13      /oracle/crs10gr2/cdata/crs10gr2
DVPORA1      2006/04/13 09:46:34      /oracle/crs10gr2/cdata/crs10gr2
```

Après avoir choisi le fichier ocr à restaurer, faire la restauration via ocrconfig -restore.

```
oracle> cd /oracle/crs10gr2/cdata/crs10gr2
oracle> sudo ocrconfig -restore week_.ocr
```

Il ne doit pas y avoir de message d'erreurs.

Redémarrage du CRS sur tous les noeuds.

```
oracle> sudo crsctl start crs

oracle> crs_stat -t

Name                Type          Target    State    Host
-----
ora...1P1.srv      application   ONLINE   ONLINE   DVPORA1
ora...1P2.srv      application   ONLINE   ONLINE   DVPORA2
```

```
ora....MB1P.cs application ONLINE ONLINE DVPPORA2
ora....P1.inst application ONLINE ONLINE DVPPORA1
ora....P2.inst application ONLINE ONLINE DVPPORA2
ora.DMB1P.db application ONLINE ONLINE DVPPORA1
ora....1P1.srv application ONLINE ONLINE DVPPORA1
ora....VP1P.cs application ONLINE ONLINE DVPPORA1
ora....P1.inst application ONLINE ONLINE DVPPORA1
ora....P2.inst application ONLINE ONLINE DVPPORA2
ora.DVP1P.db application ONLINE ONLINE DVPPORA1
ora....03.lsnr application ONLINE ONLINE DVPPORA1
ora....RA1.gsd application ONLINE ONLINE DVPPORA1
ora....RA1.ons application ONLINE ONLINE DVPPORA1
ora....RA1.vip application ONLINE ONLINE DVPPORA1
ora....06.lsnr application ONLINE ONLINE DVPPORA2
ora....RA2.gsd application ONLINE ONLINE DVPPORA2
ora....RA2.ons application ONLINE ONLINE DVPPORA2
ora....RA2.vip application ONLINE ONLINE DVPPORA2
```

Contrôles

```
oracle> crsctl check crs

CSS appears healthy
CRS appears healthy
EVM appears healthy

oracle> ocrcheck

Status of Oracle Cluster Registry is as follows :
Version                :                2
Total space (kbytes)   :           262144
Used space (kbytes)    :             5356
Available space (kbytes) :          256788
ID                     :          1891973498
Device/File Name       : /oradata/crsdvp1p/ocr_dvp1p.dbf
                       Device/File integrity check succeeded
                       Device/File not configured
Cluster registry integrity check succeeded
```

III-B - Backup de l'OCR

A toute fin utile, la commande de sauvegarde de l'OCR qu'il sera bon d'exécuter à la fin de notre réparation.

```
oracle> sudo ocrconfig -export /oradata/crsdvp1p/backup_ocr/ocr_du_29112006.dbf -s online
```

Par rapport à la place que ces fichiers prennent et au service qu'ils peuvent rendre, il est utile d'en garder des sauvegardes régulières et nombreuses en local.

IV - Procédure manuelle

Supprimer toute la configuration DVP1P de l'ocr, en forçant avec l'option -f

- -d Nom de la base
- -f Option FORCE, nécessaire dans notre cas de figure

```
oracle> srvctl remove database -d DVP1P -f
```

Réinscrire la base DVP1P dans l'ocr

- -d Nom de la base
- -o Localisation des binaires Oracle (ORACLE_HOME)
- -p Localisation du spfile central si celui-ci ne se trouve pas sous $\${ORACLE_HOME}/dbs$

```
oracle> srvctl add database -d DVP1P -o /oracle/db10gr2 -p /oradata/dbdvp1p/admin/spfileDVP1P.ora
```

Associer les instances à la base, en spécifiant sur quel noeud se trouve chaque instance

- -d Nom de la base
- -i Nom de l'instance
- -n Nom du noeud sur lequel se situe l'instance

```
oracle> srvctl add instance -d DVP1P -i DVP1P1 -n DVPPORA1
srvctl add instance -d DVP1P -i DVP1P2 -n DVPPORA2
```

Démarrer la base et les instances... la seconde instance prend un certain temps à passer en ONLINE

```
srvctl start database -d DVP1P
```

Créer le TAF , puis le démarrer. Cela peut se faire via dbca, mais la commande en ligne est spécifiée ici

- -d Nom de la base
- -s Nom du service (qui doit être différent du nom de la base)
- -r Noms des instances qui doivent être accessible pour ls service (Preferred)
- -P Type de service TAF (None, Basic ou Preselect)

```
oracle> srvctl add service -d DVP1P -s DVP1P1 -r "DVP1P1,DVP1P2" -P basic
srvctl start service -d DVP1P
```

L'état de notre CRS est maintenant d'aplomb.

```
oracle> crs_stat -t
```

Name	Type	Target	State	Host
ora...1P1.srv	application	ONLINE	ONLINE	DVPPORA1
ora...1P2.srv	application	ONLINE	ONLINE	DVPPORA2
ora...MB1P.cs	application	ONLINE	ONLINE	DVPPORA2
ora...P1.inst	application	ONLINE	ONLINE	DVPPORA1
ora...P2.inst	application	ONLINE	ONLINE	DVPPORA2
ora.DMB1P.db	application	ONLINE	ONLINE	DVPPORA1
ora...1P1.srv	application	ONLINE	ONLINE	DVPPORA1
ora...VP1P.cs	application	ONLINE	ONLINE	DVPPORA1
ora...P1.inst	application	ONLINE	ONLINE	DVPPORA1
ora...P2.inst	application	ONLINE	ONLINE	DVPPORA2
ora.DVP1P.db	application	ONLINE	ONLINE	DVPPORA1
ora...03.lsnr	application	ONLINE	ONLINE	DVPPORA1

```
ora...RA1.gsd application ONLINE ONLINE DVPPORA1
ora...RA1.ons application ONLINE ONLINE DVPPORA1
ora...RA1.vip application ONLINE ONLINE DVPPORA1
ora...06.lsnr application ONLINE ONLINE DVPPORA2
ora...RA2.gsd application ONLINE ONLINE DVPPORA2
ora...RA2.ons application ONLINE ONLINE DVPPORA2
ora...RA2.vip application ONLINE ONLINE DVPPORA2
```

DBA, tu as bien fait ton travail : va dormir en paix... pour ce qu'il te reste comme nuit...

V - Recréation complète d'un OCR corrompu

Lorsqu'il n'est plus possible de s'en sortir avec des sauvegardes d'OCR et que celui-ci semble désespérément corrompu, ne reste plus que la solution ultime (avant la réinstallation totale !) de le nettoyer totalement. Cette procédure vient de la documentatin d'Oracle-même : note Metalink 399482.1. La voici résumée: veillez à exécuter tous ces ordres de manière séquentielle et précise !

A exécuter sur tous les noeuds du RAC, en tant que root

```
${CRS_HOME}/install/rootdelete.sh
```

N'exécutez par contre pas le script `${CRS_HOME}/install/rootdeinstall.sh` qui aurait pour conséquence de supprimer les binaires CRS, vous obligeant alors à passer par une réinstallation du logiciel de clusterware.

Contrôlez ensuite que tout processus Oracle soit arrêté sur tous les noeuds. Les ordres suivants ne doivent retourner aucun processus :

```
ps -e | grep ocs
ps -e | grep crs
ps -e | grep evm
```

A exécuter sur tous les noeuds du RAC, en tant que root

```
${CRS_HOME}/install/root.sh
```

Ne reste plus qu'ensuite à repasser par le chapitre IV pour recataloguer tous les composants.