

Installing Oracle 9i release 1 RAC on Windows/2000

This technical note gives some hints and tips on installing Oracle Real Application Clusters, 9.0.1.4 on a Windows/2000 cluster. Although an HP/Compaq Proliant system with two DL380 systems in a cluster with an external MSA1000 storage array were used for the actual setup, it is likely to be relevant for other configurations as well.

This document should *not* be used as the only guidelines for installing Oracle9i RAC on Windows; you should also consult relevant Oracle and hardware documentation. A list of references is found at the end of this paper.

Preparations

Before starting the actual Oracle installation, you need some preparations:

Configuring shared disks for Oracle use

You need to make sure your external disks are configured properly for shared use by your nodes. In the case of the system named above, the disk array is delivered with a utility for configuration; you should use this to configure logical volumes for use by Oracle. In our case, the disk array was equipped with 14 physical disk drives, which we put into three logical volumes with respectively 10, 2 and 2 disk drives with RAID0+1 mirroring and striping. The large volume would serve to store all datafiles, the second one would store redo log files and the last one would store archived log files.

Subsequently, you use the Windows Disk Management utility (Start->Programs->Administrative tools->Computer Management->Disk Management) to split the volumes into the necessary logical drives for use as the raw devices serving as redo log files, control files, and the database files of all tablespaces.¹ This utility seems very happy to assign drive letters to all logical drives, we therefore needed to explicitly remove those drive letters from both nodes before continuing. Note, that the archived log files are written to ordinary file systems (NTFS), and each node needs to have its own. Again, it is imperative, that you make sure each logical drive used for this purpose only has drive letters assigned on the single node, where it will be used. Please double-check that no drive letters are assigned to raw devices (for database files, control files and redo log files) and that a drive letter is only assigned on the single node, where each NTFS formatted drive will be assigned as the archive log destination.

Configuring the interconnect

Although Oracle RAC will run without a fast interconnect between the nodes, it is highly recommended to have one. Typically you would configure the ordinary LAN (e.g. using 100Mbit Ethernet) with the public names of the nodes and the interconnect (e.g. using 1Gbit Ethernet or fibre-optics) separately with its own names. In our case, the LAN was in IP address range 192.168.10.xxx and the interconnect was configured in the IP address range 172.16.144.xxx. You also need to configure the node names

¹ Oracle now delivers a Clustered File System for use on Windows/2000. This is officially part of Oracle9i Release2, and allows you to use ordinary (clustered) files as database files rather than raw devices. Some Unix platforms also have a clustered file system available.

properly, in our case the LAN had an external name server available and we put the names for the interconnect addresses in the hosts file found in `\winnt\system32\drivers\etc`. For the interconnect, you only need TCP/IP and no Windows protocols, but (temporarily) adding Windows file sharing enables you to perform a test, where copying a large file can give you assurance, that the fast interconnect is actually used.

If you install and configure the interconnect after software install and database creation, please refer to reference [4], which also explains how you can add information about an extra node added to your cluster.

It is important to make sure each node is accessible via the default Windows shares from the other nodes. If you e.g. are on node1, you should be able to get to `\\node2\c$` and vice versa using the nodes public (LAN) names. The Oracle Installer uses this to copy files across. This can be verified by the `clustercheck.exe` utility, it is however our experience, that this utility is not working properly, and the only real verification needed is that the nodes can access each other using the default shares.

Oracle pre-install tasks

On the 9.0.1.1 CDROM, you will find a `\rac_install` directory with sub-directories. You need to find and execute the “`clustersetup.exe`” utility before actually starting the Oracle installation. This will install the so-called OSDs or Cluster Ware for RAC and will configure the nodes. When prompted, enter the private (interconnect) and public (LAN) node names of all your nodes in the cluster. If you don’t have an interconnect, use the public LAN node names in all cases. You need to do this on all nodes.

Unless you want to use the Database Creation Assistant to create your database with the Oracle default tablespace names, you don’t need to configure any logical drives as raw devices at this time, except the one referred to as the “quorum partition”.

The OSD install will create two Windows/2000 services. One, the Oracle Object Manager Service is responsible for making logical drives available to Oracle, and the other, the Oracle CM Service, is responsible for the necessary communication between the nodes. Both services will be started automatically during a system reboot. If you experience problems, e.g. on the interconnect, stopping and re-starting the Oracle CM Service may be necessary.

Oracle software installation

To install the latest release of Oracle9i, release 1, you need to have three sets of media: Oracle 9.0.1.1 software and Oracle 9.0.1.4 upgrade, and you need Oracle patch #2317441, which is available from Metalink.

Installing Oracle9i 9.0.1.1

The installation of the Oracle software is quite straight forward, in particular if you choose a “software only” install. Towards the end of the installation, you may get an error saying that a certain file in a directory called `\preinstall_rac` is not found – this directory is found on the first CD, so you need to re-insert this and press the Retry button. When the installation has completed, the entire installation will be copied to

the other node(s); this is rather time consuming, and the only indication of progress is to actually inspect the Oracle directories on the other node(s).

Upgrading to 9.0.1.4

The upgrade to 9.0.1.4 is straightforward and uses the Oracle installer.

Installing Patch 2317441

The readme file included with this patch specifically says, it should be installed on a 9.0.1.3.1 release only. However, you also need it for 9.0.1.4 as it includes important changes to the OSD's (Cluster Ware) that are not part of the 9.0.1.4 patch. The simplest way to install patch 2317441 is to copy the following list of files directly from the preinstall_rac\osd directory on the patch distribution to the respective locations in \winnt\system32\osd9i:

```
osd\CM.dll  
osd\CMSrvr.exe  
osd\OraFencedrv.sys  
osd\ipc_tcp.dll  
osd\ipc_via.dll  
osd\mscs\ORACAvail.dll  
osd\mscs\ORACAvailEx.dll  
osd\start.dll
```

This must be done on all nodes and should preferably be done before any database is created.

Database creation

Creating a database for RAC is somewhat more complex than a single-instance database. Chapter 5 in reference [6] provides detailed information about database creation, and chapter 2 in reference [5] provides detailed information about init.ora parameters.

Preparing logical drives as raw partitions

Oracle RAC needs raw partitions¹ as database files, and these need to be configured using the Oracle utility called Object Link Manager. It is found in \winnt\system32\osd9i\olm, where the GUIOracleOBJManager.exe utility will bring up a screen allowing you to configure symbolic link names for all your logical drives. We recommend creating a shortcut for this utility, e.g. on your desktop. The utility will bring up a list of all configured logical drives, and will allow you to name these for use by Oracle. After adding names, choose 'commit' from the menu, and you are done. If you are using your own create database/create tablespace scripts, you can choose any naming scheme you want, as long as it follows the \\.\file_name standard. Configure all database files and redo log files using this. The setup is automatically duplicated to other nodes.

Creating the database

The distinction between an instance running as shared and exclusive is done via the `cluster_database` parameter in your `init.ora` file². During database creation, use `cluster_database=false`.

In addition to running the catalog scripts that are part of your normal database creation (such as `catalog.sql` and `catproc.sql`) you need to run `catclust.sql`.

You need to run the ORADIM utility on both (all) nodes, giving the instances different system identifiers, e.g. `SID1` and `SID2`. Also remember the `thread=N` and `instance_number=N` parameters in all instances `init.ora` files. If you use rollback segments, you need separate segments for each instance; if you use system managed undo, you need an undo tablespace for each instance. Similarly, each instance needs to have its own set of redo log files (which have to be shared as raw devices). Chapter 5 of reference [6] discusses the details of this.

After the database has been created, you can change to `cluster_database=true` and start an instance on both (all) nodes. Except for the parameters `thread`, `instance_number` and `rollback_segments` (or `undo_tablespace`), the `init.ora` files of all instances can be identical. If you want to use SPFILE rather than traditional `init.ora` files, the SPFILE must refer to a single shared raw device.

Post installation

There are a few activities that you would run after installation and after verification of your configuration.

Turning on archive log mode

To turn on archive log mode, you need to bring down both (all) instances, change to `cluster_database=false`, bring up one instance and mount but not open the database and change it to archive log mode. Subsequently, you can shut down the instance, change back to `cluster_database=true` and bring up both (all) instances. The archive log destination of the instances must be different, although you must make sure both (all) archive log destinations can be made available to one node in case you need to run a recovery on a single instance.

Backup of a cluster database

Your backup procedure needs to correctly handle both the raw devices and the dual (multiple) locations of archived log files.

Running statspack

As statspack queries instance specific data you need to have statspack data collection taking place on all instances. Install statspack the way you would normally install it, typically using a `perfstat` user and a specific tablespace, and start a job to run at regular intervals (e.g. every half or full hour) on all instances. Note, that you should specify the instance parameter to the `dbms_job.submit` procedure.

² In Oracle8 Oracle Parallel Server you did 'startup shared' or 'startup exclusive' – although this syntax still is in place it has no effect.

Configuring transparent failover

If you need to configure SQL*Net to automatically fail over from one instance to the other in case of an instance failure, you need to specify failover_mode in your tnsnames.ora file. A sample entry is:

```
oradb.company.dk =
  (description =
    (address_list=
      (load_balance=on)
      (failover=on)
      (address=(protocol=tcp)(host=192.168.10.1)(port=1521))
      (address=(protocol=tcp)(host=192.168.10.2)(port=1521))
    )
    (connect_data=
      (service_name=oradb)
      (failover_mode=(type=select)(method=basic)(backup=oradb.company.dk))
    )
  )
```

The failover=on entry means connection will be attempted until a working system is found, load_balance=on means connection will chose randomly between the systems. The failover_mode entry tells how transparent application failover is done and which tnsnames.ora entry is used; this effectively points to its own entry. Please note, that the backup=... entry must be present, despite the fact that some examples in chapter 9 of [6] does not show this.

References

- [1] Step-By-Step Installation of RAC on Windows NT/2000, Metalink document #178882.1
- [2] Oracle9i Database Installation Guide for Windows, which has an appendix on Real Application Clusters should be read as well.
- [3] Compaq Parallel Database Cluster Model PDC/O2000 for Oracle9i and MSA1000 on Windows 2000, Compaq Part No. 274734-001
- [4] How To Change The Oracle Cluster Manager Settings on Windows, Metalink document #184840.1
- [5] Oracle9i Real Application Clusters Administration, Oracle Part No. A89869-01
- [6] Oracle9i Real Application Clusters Installation and Configuration, Oracle Part No. A89868-01

Revisions

Release 1: 26-Aug-2002.

Release 2: 17-Jan-2003 updated for release 9.0.1.4

Copyright ©, 2002 by Miracle A/S