

Adding a Node to 10gR1 on RHEL4

By Jeff Needham

Trying to add a node, to demonstrate how NON dynamic provisioning works for RAC.

Background:

The installer needs to 'rsh' or 'ssh' to any of the other nodes without any Interaction of messages, passwords etc. This is used to detect if the nodes are there and to copy files to them during the install.

If you cocked it up, then the installer fails without understanding what has gone wrong. It just says: "I can't get to the node". And really gives no reason or diagnostics as to what has gone wrong.

In fact the Installer in general is the most non-adaptive piece of software Oracle every attempted. It is very brittle and at the first sign of trouble, it fails and really doesn't provide a good sense of diagnosis in the process.

I don't know how many of you are attempting this, but even if you didn't have a missing libstdc++.so, this is pretty convoluted. It is pretty noisy and unedited (much like me :-)

However I am now running my 5 node 10gR1 cluster on the R2 clusterware (which is supported and recommended). This is a 64-bit install on RHEL4.2 (so a little edgy)

Go get Metalink note (270512.1)

1. If the CRS Home is owned by root and you are on a version < 10.1.0.4, change the ownership of the CRS Home directories on all nodes to the Oracle user so that OUI can read and write to these directories.

I'm on 104 64-bit

Connections, backbone, iSCSI fabric all the same.

This really is just add node 5 to the CRS bit.

The /etc/sysconfig/rawdevices needs to be configured ... Started ... Permissions changed. OK good.

2. Set the DISPLAY environment variable and run the addNode.sh script from \$ORA_CRS_HOME/oui/bin on one of the existing nodes as the oracle user.
Example:

```
DISPLAY=ipaddress:0.0; export DISPLAY
cd $ORA_CRS_HOME/oui/bin
./addNode.sh
```

In this case, I have installed, mapped and mounted GFS (including /ohome which is

where the kernel is) CRS is locally on /home/jneedham on all nodes (no shared CRS home, bucko)

I only have one other node up (1) at the moment. (this might be a problem)

Fire up X on my Mac (Panther 10.8 X, so it behaves pretty well)

Run xhost + on the Mac

Test with xlcock ... good

Running addNode.sh from Node1 (subbu1)

Checking operating system version: must be UnitedLinux-1.0, redhat-3 or SuSE-9

Failed <<<<

Grrr. Forgot to hack the /home/jneedham/1010/product/crs/oui/oraparm.ini (this is still R1 CRS) to add "redhat-4" to the Certified-Versions list.

More Grrr. Forgot to do all the keys stuff, which also requires you to ssh to everywhere to get rid of the "... RSA key fingerprint is ca: ..." interactive question when stuff is being added to .ssh/known_hosts

Oh no. I have to bring up all the nodes in the cluster

Now it is complaining that remote nodes don't have writeable homes

I do have to change permissions it seems ... Unlike the note above

OK, that worked.

I just got to the Node Addition Summary page and am about to do the install

Copying stuff to new node In progress ...

OK, it is asking me to run rootaddnode.sh on subbu1 (which is node 1)

Hmmm OK, I'll bite

clscfg: EXISTING configuration version 2 detected.

clscfg: version 2 is 10G Release 1.

Attempting to add 1 new nodes to the configuration

Using ports: CSS=49895 CRS=49896 EVMC=49898 and EVMR=49897.

node <nodenumber>: <nodename> <private interconnect name> <hostname>

node 5: subbu5 subbu5_gcs subbu5

Creating OCR keys for user 'root', privgrp 'root'..

Operation successful.

Ok, I'm off to node5 to run root.sh

Checking to see if Oracle CRS stack is already up...

/etc/oracle does not exist. Creating it now.

/bin/cp: cannot create regular file '/dev/raw/raw1': Device or resource busy

Huh? Owned by the right person,

I'm going to reboot subbu5 and see if that helps.

I have no idea why I can't do a cp /dev/null to /dev/raw/raw1.

*I reran root.sh after I rebooted and that worked better.
I had brought up rawdevices on this box on the fly (not from a dead stop), so
I just rebooted and root.sh is working better. The installer on node1 is still running
Waiting for an OK.
Got a LOT further.
Now it is hung at startcheck ‘-(
I aborted it again, so I don’t know if I have a node5 yet.*

I tried to run init.crs on node 5 and got this:

```
/home/jneedham/1010/product/crs/bin/crsd.bin: error while loading shared libraries:  
libstdc++.so.5: cannot open shared object file: No such file or directory
```

*This is fried! It turns out, this specific .so was never loaded on this node, despite the
install being just like the other nodes.*

*I wiped everything and did a scratch 5 node CRS install. When I came to node 5 this
died again. I manually moved libstdc++.5.0.7 and created this link and everything
was fine. This was the ONLY thing that was missing. Weird.*

*So, basically, it didn’t work, but might have worked if libstdc++.so.5 was actually
installed by the OS.*

*This is what makes the selling feature of adding CFS nodes (from GFS, OCFS or
Matrix) easily a non-selling point for RAC.*

I have to go in and completely debug this from scratch.

I usually just blow it all away and start from scratch and now I know why I do that.

*One thing that should be completely STRESSED is that If you forgot to ssh to all the
nodes in the combination to get rid of the message about adding hosts to
.ssh/known_hosts, then the thing will fail.*

*Not only does it fail, it doesn't really know what went wrong, let alone how to fix it.
If you have ever done the ssh_keygen -d dsa, copy the crap around then then ssh until
you are blue in the face. Sounds like a good job for a computer!*

One Final Comment from Bruce:

So yes -- its whack'd -- very whack'd

*Just to mess with your head for a while .. really adding a node is adding a node-named
key to a slot in the voting disk -- clscfg -addnode <edited>. And then vreating a stub
for your resources racgons -addnode in the OCR .. after that its gravy .. the addnode
script itself is buggy as hell -- take note that the stupid thing over-writes the \$HOME
env var and then everything sucks. Its only purpose in life is to copy the
\$ORA_CRS_HOME and ORACLE_HOME to a new node and set up the inventory.*

*So if one wanted to (goofball like me perhaps) copy over the homes and oraInventory
set to the new node and do the clscfg and racgons -- life is good and sucks less.*

And to mess with your head a bit more .. think about what happens if some bonehead wipes your OCR and Voting disk on NAS or OCFS and you don't have a backup. You can restore your OCR from cdata but there does not exist a procedure to re-create the voting disk. Without it the cluster will flop around on the floor rebooting itself. Here is the trick... save a copy of your OCR if you have it .. then remove the OCR and Voting disk files -- run \$CRS_HOME/root.sh (edited with the right info indeed) on your #1 node and viola a blank OCR and registered voting disk are created. -- then treat the cluster as a new install and run root.sh on the nodes to bring up CRS - After that you can restore your OCR from cdata or previous save and you are in business.

Lastly for kicks : take a scan of a dd'd copy of your voting disk .. if you add node and delete node a few times you will see that the SKF piece of <edited> collects junk like velcro -- if you delete a node (properly) the node keys and data are left behind in Voting disk and OCR.