



NYOUG

New York Oracle User Group
Fall 2012 General Meeting



DBA for 25 years, 17 yrs using Oracle.

UNIX (Solaris, Aix, Windows, Linux)

Multi-year project to migrate to RAC One-Node.

Chemical manufacturing, publishing, retail, finance.

10g OCA

Contact – [john.larkin1 @comcast.net](mailto:john.larkin1@comcast.net)



Reduce costs and

Replace Third-Party Cluster Services

with Grid Infrastructure and RAC One-Node

(Infrastructure Technologies Track)



- **Vendors of clustering software.**
- **Oracle's RAC One-Node – integrated, single vendor, opportunity for reduced cost.**
- **Basic Grid Infrastructure required for RAC One-Node.**
- **Multi-year project to migrate 140+ databases.**
- **High-level survey of the functionality of Oracle RAC.**
- **Cost estimates for RAC One-Node vs. RAC.**



- **Explore basic concepts and tools.**
- **11g R2 Oracle database / Grid Infrastructure cluster.**
- **ASMLib driver, Linux RPM's, Filesystem layout,**
- **Grid Infrastructure installation and a patching workaround, Oracle RDBMS installation,**
- **Oracle RAC One-Node Database installation,**
- **RAC One-Node relocation from one node to another.**
- **A 2-Node cluster with a RAC One-Node database.**
- **Out Of Scope:**
 - **Cluster support for Applications (simulate RAC 1)**



- **IBM(HACMP) / Microsoft(Clustering Service) –**
 - shared-nothing
- **Veritas/Symantec(VCS)**
 - shared-disk.
- **Oracle (Grid Infrastructure)**
 - shared-disk.
- **Grid Infrastructure:**
 - inter-instance communication via interconnect.
 - **Cache Fusion – global buffer cache**
 - In-memory data block changes available to all



HP Blade servers x86-64, 8 CPU, 48G Ram

- **3 - HP Blade servers x86-64, 8 CPU, 48G Ram**
 - **Dataguard for some – 3 additional Blades.**
- **6 - GigE NIC's**
- **RedHat Enterprise Linux Release 5.3**
- **Oracle 11g R2.02**



- **Do the Up-front work, save time and effort.**
- **Network Setup**
 - **VIP names**
 - **IP addresses**
 - **SCAN names**
 - **IP addresses**
 - **Round-robin setup**
 - **Interconnect**
- **DB Ports**



- **Backbone of Oracle's clustering capabilities**
 - **HA services monitor the cluster**
 - **Oracle Database / RDBMS binaries**
 - **Oracle Grid Infrastructure (Clusterware) binaries.**
- **Grid Infrastructure ORACLE_HOME – "GRID_HOME"**
- **Database ORACLE_HOME or "RDBMS_HOME".**
- **Tightly coupled with ASM**
- **Fully integrated in 11g R2**



- **Split personality :**
 - **non-RAC when static (not in a state of flux)**
 - **full RAC database when relocating**
 - **Multiple UNDO tablespaces**
 - **Multiple threads**



- **Split personality : (cont'd)**
 - **Query gv\$thread - inst_id and status columns**
 - **One-Node only see INST_ID=1**
 - **1 STATUS=OPEN, 1 - STATUS=CLOSED**
 - **Full RAC - see more than 1 INST_ID, all STATUS=OPEN.**



- **SQL> select * from gv\$thread;**

INST_ID	THREAD#	STATUS	ENABLED	GRP	INSTNCE
1	1	OPEN	PUBLIC	2	PCT2Q_1
1	2	CLOSED	PUBLIC	2	PCT2Q_2



```
SQL> select * from gv$thread;
```

INST_ID	THREAD#	STATUS	ENABLED	GRP	INSTNCE
1	1	OPEN	PUBLIC	2	BRNYT1
1	2	OPEN	PUBLIC	2	BRNYT2
1	3	OPEN	PUBLIC	2	BRNYT3
2	1	OPEN	PUBLIC	2	BRNYT1
2	2	OPEN	PUBLIC	2	BRNYT2
2	3	OPEN	PUBLIC	2	BRNYT3



```
SQL> select tablespace_name, contents, status  
       from dba_tablespaces;
```

TABLESPACE_NAME	CONTENTS	STATUS
-----	-----	-----
ORACLE_TRN_UNDO_1	UNDO	ONLINE
ORACLE_TRN_UNDO_2	UNDO	ONLINE

...

(one tablespace for each instance)



```
select comp_id, status, version, comp_name from
  dba_registry;
```

CompID	STATUS	Version	CompNm	INSTNCE
CATALOG	VALID	11.2.0.2.0	ODb Catalog Views	
CATPROC	VALID	11.2.0.2.0	ODb Pkgs + Types	
JAVAVM	VALID	11.2.0.2.0	JServer JAVA VM	
CATJAVA	VALID	11.2.0.2.0	OraDb Java Pkgs	
RAC	VALID	11.2.0.2.0	Oracle Real Application Clusters	



- **Network Requirements**
 - **Same as a Full RAC database**
 - **Typically 3 NICs min.**
 - **Private – Interconnect**
 - **Public – SCAN, VIP's**
 - **I/O – Network-attached storage**



- **SCAN (Single Client Access Name) - single name**
- **DNS - resolve to 3 IP addresses - round-robin**
- **SCAN entries to the /etc/hosts file as comments**
 - **(temp workaround - /etc/hosts with a single IP).**
- **SCAN resources :**
 - **SCAN Vip**
 - **SCAN Listener**
 - **for each IP defined in the DNS SCAN entry.**
- **SCAN Listener runs from GI_Home.**



- **ASM – Automatic Storage Management**
- **AsmLib (ASM Library driver) RPM**
 - **highly dependent on the OS kernel version.**
 - **rpm -qa --queryformat "%{NAME}-%{VERSION}-%{RELEASE} (%{ARCH})\n" | grep oracleasm | sort**
 - **Error log not verbose**
 - **Administer using /usr/sbin/oracleasm**
 - **/var/log/oracleasm**
 - **I was bitten by this during an OS patch**



```
echo "( asm raw dev in /dev/mapper/ OR dm- in /dev/)"
for i in `/sbin/multipath -v 1 -l | grep "^asm" | sort` ; do
  v_asmdisk=$i
  v_size=`/sbin/multipath -ll $v_asmdisk | awk -F[ '{print
    $2}'|awk -F] '{print $1}' | grep size` ; # lines w size
  v_device=`ls -alF /dev/mpath/$v_asmdisk|awk '{ print
    $11 }' |awk -F"." '{print $3 }'|awk -F"/" '{print $2}` ;
  echo $v_asmdisk $v_size "/dev/"$v_device |awk '{
    print "ASM candidate "$1 " " $2 " " $3 }';
done }
```



(asm raw device in /dev/mapper/ OR dm- in /dev/)

ASM candidate asmocr01	size=10G	/dev/dm-5
ASM candidate asmsys01	size=10G	/dev/dm-10
ASM candidate asm1p001	size=30G	/dev/dm-18
ASM candidate asm1par1	size=30G	/dev/dm-13
ASM candidate asmacf01	size=10G	/dev/dm-8
ASM candidate asmbu01	size=125G	/dev/dm-25



- **The basic oracleasm commands are:**
(arranged in order of use)
 - **init** **Load & initialize ASMLib driver**
 - **configure** **Configure Linux ASMLib driver**
 - **createdisk** **Allocate device for Oracle's use**
 - **scandisks** **Scan for known ASMLib disks**
 - **status** **Display status of ASMLib driver**
 - **listdisks** **List known ASMLib disks**
 - **querydisk** **Display status of a specific disk**



- **oracleasm commands (cont'd) :**
 - **exit** Stop the ASMLib driver
 - **deletedisk** Return a device to the OS
 - **renamedisk** Change label of an ASMLib disk
 - **update-driver** Download latest ASMLib driver



Screen Name

Response

=====

=====

- Download S/w Updates
- Installation Option

remain UNINFORMED !!
 " Install and Configure
 Grid Infrastructure for a
 Cluster"

- Storage Option Information
 (9 of 16)

"Automatic Storage
 Management (ASM)".

- **INSTALL**

STOP at root.sh prompt



- On ALL NODES : Run OPatch
- `${GRID_HOME}/OPatch/opatch napply -oh
${GRID_HOME} -local ${MEDIA_HOME}/12419353 -
invPtrLoc
/oracle/grid_base/orainventory/orainst.loc`
- `${GRID_HOME}/OPatch/opatch napply -oh
${GRID_HOME} -local ${MEDIA_HOME}/12419331 -
invPtrLoc
/oracle/grid_base/orainventory/orainst.loc`



- As the GRID user in the GRID_HOME :
 - asmca &
 - Disk Groups "Disk Groups" tab,
Create" button.
 - Create Disk Group see 2 ASMLib vol
"Change Disk Discovery
Path"
 - "Redundancy" " External (none)" except
OCRDG01. Don't select
QUORUM. (voting disks)



- As ORACLE user into ORACLE_HOME
 - From /oracle/Oracle11gR2.11.2.0.2/database
 - ./runInstaller &

Screen Name

Response

=====

=====

- | | |
|-----------------------|--------------------------------------------------|
| • Installation Option | "Install db software only". |
| • Grid Options | "Real Application Cluster Database installation" |
| • INSTALL | STOP at root.sh prompt |
| • PATCH ORACLE_HOME | |



- As ORACLE user in ORACLE_HOME
 - dbca &

Screen Name

Response

=====

=====

- Welcome Screen
- Db File Loc
- Database Area:

RAC One Node database.
 Storage: Auto Stg Mgmt (ASM)
 +ORACRS



- **srvctl config listener**
Name: LISTENER
Network: 1, Owner: grid
Home: <CRS home>
End points: TCP:1521
- **srvctl status listener**
Listener LISTENER is enabled
Listener LISTENER is running on node(s):
pctcr10, pctcr11



- **srvctl modify listener -l LISTENER -p "TCP:1760"**
- **srvctl stop database -d pct002d**
- **srvctl stop listener -l LISTENER -n pctcr10**
- **srvctl start database -d pct002d**
- **Listener should autostart**

srvctl config listener

Name: LISTENER

Network: 1, Owner: grid

Home: <CRS home>

End points: TCP:1760



- Relocate to alternate node. (As GRID owner)
 - Server overloaded, O/S maintenance
- NODE1 :
 - `srvctl status database -d pct002t`
Instance PCT002T_1 is running on node pctcr10
Online relocation: **INACTIVE** <= relocate inactive
 - `srvctl relocate database -d PCT002T -n pctcrs11`



- **NODE1 :**
- **srvctl status database -d pct002t**
Instance PCT002T_1 is running on node pctcrs10
Online relocation: **ACTIVE** <= relocation active
Source instance: PCT002T_1 on node pctcrs10
Destination inst: PCT002T_2 on node pctcrs11



- **NODE2:**
- **srvctl status database -d pct002t**
Instance PCT002T_2 is running on node pctcrs11
Online relocation: **INACTIVE** <= reloc'n inactive
- **Ps -ef |grep pmon**
oracle 2544 10:06 00:00:00 ora_pmon_PCT002T_2
grid 28488 Jan31 00:00:00 asm_pmon_+ASM2



- export ORACLE_SID=PCT002T_2
- Sqlplus / as sysdba

SQL> show parameter undo

NAME	VALUE
undo_management	AUTO
undo_retention	900
undo_tablespace	UNDOTBS2 <<==*



- **If time permits.**
- **KFED- good info in dumps**



- **Handy utilities**
 - **asmmap** **_inactive** (asmmap similar)
 - **oracleasm**
 - **srvctl**
- **A Cheat for Patching**
- **Temporary workaround for network issues**
- **Gotcha's**
 - **ASMLib – OS version dependence**
 - **Local_listener – 1521**



- **Many features available in the Oracle Clusterware**
- **Clustered database - options to reduce downtime.**
 - **maintenance will necessitate a server outage.**
 - **automated method to restart a database.**
 - **hardware problems**
 - **allow queries to be restarted after relocation.**
 - **continue to enhance the capabilities**
- **An overview of the RAC installation process**
- **Basic components**



- **Reference Material**
- Oracle 11g Administrators guide
- MOSC – Bug reports, Technical Notes, How-To articles
- Jeffrey Hunter – Articles on building a RAC system for less than \$2000.
- Vincent Chan – Moving to RAC One-Node



- **Cost considerations: RAC One-Node vs Full RAC**
- Approximately 55% less for RAC One-Node
- \$23k vs. \$10k
 - Prices are list prices – discounts should apply
- License limitation for One-Node
 - **“the most used technology in 20 yrs.... And the most oversold”.**
 - **RAC db on single server.**



THANK YOU !

Thanks to all members of our DBA team for the group effort required to make our RAC installations robust and flexible. Special recognition for Devesh and Pradeep for recommendations and insight into their experiences.



- **QUESTIONS ?**



- **Thank You !**
- **Please complete session evaluations**
 - **Session 808 – RAC One-Node**
 - **John Larkin – [john.larkin1 @comcast.net](mailto:john.larkin1@comcast.net)**