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
Reduce costs and

Replace Third-Party Cluster Services

with Grid Infrastructure and RAC One-Node

(Infrastructure Technologies Track)
Agenda

• 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
Hardware and Software

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;

<table>
<thead>
<tr>
<th>INST_ID</th>
<th>THREAD#</th>
<th>STATUS</th>
<th>ENABLED</th>
<th>GRP</th>
<th>INSTNCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>OPEN</td>
<td>PUBLIC</td>
<td>2</td>
<td>PCT2Q_1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>CLOSED</td>
<td>PUBLIC</td>
<td>2</td>
<td>PCT2Q_2</td>
</tr>
</tbody>
</table>
SQL> select * from gv$thread;

<table>
<thead>
<tr>
<th>INST_ID</th>
<th>THREAD#</th>
<th>STATUS</th>
<th>ENABLED</th>
<th>GRP</th>
<th>INSTNCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>OPEN</td>
<td>PUBLIC</td>
<td>2</td>
<td>BRNYT1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>OPEN</td>
<td>PUBLIC</td>
<td>2</td>
<td>BRNYT2</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>OPEN</td>
<td>PUBLIC</td>
<td>2</td>
<td>BRNYT3</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>OPEN</td>
<td>PUBLIC</td>
<td>2</td>
<td>BRNYT1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>OPEN</td>
<td>PUBLIC</td>
<td>2</td>
<td>BRNYT2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>OPEN</td>
<td>PUBLIC</td>
<td>2</td>
<td>BRNYT3</td>
</tr>
</tbody>
</table>

**Split Personality – Full RAC**
SQL> select tablespace_name, contents, status from dba_tablespaces;

<table>
<thead>
<tr>
<th>TABLESPACE_NAME</th>
<th>CONTENTS</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORACLE_TRN_UNDO_1</td>
<td>UNDO</td>
<td>ONLINE</td>
</tr>
<tr>
<td>ORACLE_TRN_UNDO_2</td>
<td>UNDO</td>
<td>ONLINE</td>
</tr>
</tbody>
</table>

... (one tablespace for each instance)
Identify RAC - REGISTRY

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

<table>
<thead>
<tr>
<th>CompID</th>
<th>STATUS</th>
<th>Version</th>
<th>CompNm</th>
</tr>
</thead>
<tbody>
<tr>
<td>CATALOG</td>
<td>VALID</td>
<td>11.2.0.2.0</td>
<td>ODb Catalog Views</td>
</tr>
<tr>
<td>CATPROC</td>
<td>VALID</td>
<td>11.2.0.2.0</td>
<td>ODb Pkgs + Types</td>
</tr>
<tr>
<td>JAVAVM</td>
<td>VALID</td>
<td>11.2.0.2.0</td>
<td>JServer JAVA VM</td>
</tr>
<tr>
<td>CATJAVA</td>
<td>VALID</td>
<td>11.2.0.2.0</td>
<td>OraDb Java Pkgs</td>
</tr>
<tr>
<td>RAC</td>
<td>VALID</td>
<td>11.2.0.2.0</td>
<td>Oracle Real Application Clusters</td>
</tr>
</tbody>
</table>
Network Requirements

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

- 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
ASM - multipath

- multipaths {
  multipath {
    wwid 99999999999999999999999999999999
    alias asmocr01
  }
}

- Easy naming
- Guarantee consistent assignment
ASM – asmmmap_inactive

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 $1}' | grep size`;
  v_device=`ls -a!F /dev/mpath/$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/)

<table>
<thead>
<tr>
<th>ASM candidate</th>
<th>size</th>
<th>path</th>
</tr>
</thead>
<tbody>
<tr>
<td>asmocr01</td>
<td>10G</td>
<td>/dev/dm-5</td>
</tr>
<tr>
<td>asmsys01</td>
<td>10G</td>
<td>/dev/dm-10</td>
</tr>
<tr>
<td>asm1p001</td>
<td>30G</td>
<td>/dev/dm-18</td>
</tr>
<tr>
<td>asm1par1</td>
<td>30G</td>
<td>/dev/dm-13</td>
</tr>
<tr>
<td>asmacf01</td>
<td>10G</td>
<td>/dev/dm-8</td>
</tr>
<tr>
<td>asmbu01</td>
<td>125G</td>
<td>/dev/dm-25</td>
</tr>
</tbody>
</table>
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

- 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  
  remain UNINFORMED !!

• Installation Option
  "Install and Configure Grid Infrastructure for a Cluster"

• Storage Option Information  "Automatic Storage Management (ASM)"
  (9 of 16)

• INSTALL
  STOP at root.sh prompt
Install GI - PATCH

• 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 Discovery Path"
  - "Redundancy"
  - "External (none)" except OCRDG01. Don’t select QUORUM. (voting disks)
Install RDBMS Binaries

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

<table>
<thead>
<tr>
<th>Screen Name</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Installation Option</td>
<td>&quot;Install db software only&quot;.</td>
</tr>
<tr>
<td>Grid Options</td>
<td>&quot;Real Application Cluster Database installation“</td>
</tr>
<tr>
<td>INSTALL</td>
<td>STOP at root.sh prompt</td>
</tr>
<tr>
<td>PATCH ORACLE_HOME</td>
<td></td>
</tr>
</tbody>
</table>

27
Create RAC One-Node DB

- As ORACLE user in ORACLE_HOME
  - `dbca` &

<table>
<thead>
<tr>
<th>Screen Name</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welcome Screen</td>
<td>RAC One Node database.</td>
</tr>
<tr>
<td>Db File Loc</td>
<td>Storage: Auto Stg Mgmt (ASM)</td>
</tr>
<tr>
<td>Database Area:</td>
<td>+ORACRS</td>
</tr>
</tbody>
</table>
• Feature of 11.2 OUI:
  • does not modify the local_listener port.
  • Security issue to have default port# (1521)
  • local_listener
    (DESCRIPTION=(ADDRESS_LIST=(ADDRESS=(PROTOCOL=TCP)(HOST=xrac1-vip1)(PORT=1521))))

• Modify cluster as GRID user (GRID_HOME)
• MOSC - Change default port : Doc ID: 359277.1
  NOTE:99721.1
• `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
Modify RAC Cluster Settings

- `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

<table>
<thead>
<tr>
<th>NAME</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>undo_management</td>
<td>AUTO</td>
</tr>
<tr>
<td>undo_retention</td>
<td>900</td>
</tr>
<tr>
<td>undo_tablespace</td>
<td>UNDOTBS2 &lt;&lt;==*</td>
</tr>
</tbody>
</table>
DISKGROUPS for ACFS

• 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.

• 10-day rule – calendar year.
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