S317077: Lessons from the RAC Pack: Oracle Real Application Clusters on Oracle VM - Best Practices

Saar Maoz
RACPack – Oracle RAC Development

Roger Lopez
Oracle Database Solutions - DELL
The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle’s products remains at the sole discretion of Oracle.
Agenda

- Oracle RAC & Oracle VM Overview
- Oracle RAC on Oracle VM Best Practices
- Oracle RAC – Oracle VM Templates
- Partner Perspective – Dell
- Demos
  - Standard Build
  - Dom0 Build
  - Add / Remove Nodes / Instances
  - Live Migration
Oracle RAC Architecture

Node 1
- **VIP1**
- Service
- Listener
- instance 1
- ASM
- Managed by ASM
- RAW / Block Devices
- Oracle Clusterware
- Operating System

Node 2
- **VIP2**
- Service
- Listener
- instance 2
- ASM
- Managed by ASM
- RAW / Block Devices
- Oracle Clusterware
- Operating System

Node n
- **VIPn**
- Service
- Listener
- instance n
- ASM
- Managed by ASM
- RAW / Block Devices
- Oracle Clusterware
- Operating System

Public network

- Redo / Archive logs all instances
- Database / Control files
- OCR and Voting Disks

Managed by ASM

 RAW / Block Devices
Oracle RAC in Oracle VM Architecture
RAC on OVM – Deployment Configuration

• Two types of deployments
  – Production Oracle RAC OVM Configurations [Supported]
    • Each RAC node must be deployed on a separate physical server for production environments
    • Shared database disks must be on physical disks
  – Test Oracle RAC OVM Configurations
    • Above requirements are relaxed
Oracle RAC on Oracle VM

Production Deployment

Shared Cache / RAC

DOM-0
Guest VM

Hypervisor

X86/64 Bare-Metal Server

Dom0 OS

Guest VM

Operating System

Oracle Clusterware

Oracle ASM

Listener

Service

instance 1

instance 2
Oracle RAC on Oracle VM
Non-Production deployment
RAC OVM Templates - Availability

• Available for 11.1.0.7.2, 11.2.0.1.2, 11.2.0.2.0 on Linux 32 and 64 bits
  – Download from e-delivery or My Oracle Support
    Note:1185244.1: https://support.oracle.com/oip/faces/secure/km/DocumentDisplay.jsp?id=1185244.1

• Templates are distributed as archive files containing two disk images
  – Oracle Enterprise Linux 5 U4 system disk image
  – Oracle RAC install disk image (Clusterware, Database, ASM)
    • All homes updated to latest Bundle / CPU Patch level

• Entire install is automated, with the user only providing minimal input parameters.
RAC OVM Templates - Delivery Mechanism

- 32 bit and 64bit versions
- Image files are built with ‘sparse file’ support
- Inside each zip are tgz archives with the following files:
  - VM Config file (text file)
  - Opatch lsinventory sample output
  - README / PDFs for installing
  - Disk 1 image file: Operating System
  - Disk 2 image file: Oracle Software
    - Includes Clusterware, ASM and RAC
RAC OVM Templates - Features

• Builds production ready Oracle RAC (Cluster) in about 30 minutes
  – 11.1: 2-node, then manual add node
  – 11.2: N-node initial build, automated add/remove nodes

• 11.2 Templates allow
  – Standard 2-node build (Interview based)
  – Automatic network setup directly from Dom0
  – N-node build, also directly from Dom0
  – Run as root, oracle/grid users (w/sudo configured on build node only)
  – Automatic add/remove any number of nodes or instances
  – Options to modify environment after a build (Create DB, add DB Console)
  – Optionally create an ACFS filesystem on all nodes
  – Support for role separation or non-role separation (default)
Simple 2 Node Interview

<table>
<thead>
<tr>
<th>NODE DETAILS</th>
<th>NODE 1</th>
<th>NODE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Name</td>
<td>[test170]</td>
<td>[test171]</td>
</tr>
<tr>
<td>Public IP</td>
<td>[192.168.1.170]</td>
<td>[192.168.1.171]</td>
</tr>
<tr>
<td>Private Name</td>
<td>[test170-priv]</td>
<td>[test171-priv]</td>
</tr>
<tr>
<td>Private IP</td>
<td>[10.10.10.170]</td>
<td>[10.10.10.171]</td>
</tr>
<tr>
<td>VIP Nodename</td>
<td>[test170-vip]</td>
<td>[test171-vip]</td>
</tr>
<tr>
<td>VIP IP</td>
<td>[192.168.1.172]</td>
<td>[192.168.1.173]</td>
</tr>
</tbody>
</table>

GLOBAL DETAILS

<table>
<thead>
<tr>
<th>Domain Name</th>
<th>[localdomain]</th>
<th>DNS Server IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Network Adapter Name</td>
<td>[eth0]</td>
<td></td>
</tr>
<tr>
<td>Public Subnet mask</td>
<td>[255.255.255.0]</td>
<td></td>
</tr>
<tr>
<td>Default Gateway IP</td>
<td>[192.168.1.1]</td>
<td></td>
</tr>
<tr>
<td>Private Network Adapter Name</td>
<td>[eth1]</td>
<td></td>
</tr>
<tr>
<td>Private Subnet mask</td>
<td>[255.255.252.0]</td>
<td></td>
</tr>
</tbody>
</table>

CLUSTER DETAILS

Enter the Cluster Name: [jun16-17071] | [scan-170171] [192.168.1.174]

Do you want to configure this cluster? YES/NO: >YES<
Partner Perspective – DELL
Building Clusters with more than 2 nodes

True Silent Install

- Customer supplies an initialisation file. (netparams.ini)
- Stamp file into shared storage
- Repeat this section, identifying the 6 attributes for each node
- Power on new nodes and pass command on ‘grub’ boot-up line

**netparams.ini**

```ini
# Node specific information
NODE1=test170
NODE1IP=192.168.1.170
NODE1PRIV=test170-priv
NODE1PRIVIP=10.10.10.170
NODE1VIP=test171-vip
NODE1VIPIP=192.168.1.172

NODE2=test171
NODE2IP=192.168.1.171
NODE2PRIV=test171-priv
NODE2PRIVIP=10.10.10.171
NODE2VIP=test171-vip
NODE2VIPIP=192.168.1.173
```
Oracle RAC on Oracle VM Best Practices

- **Oracle RAC in Oracle VM whitepaper on OTN**

- **Oracle VM Server**
  - Size as you would a bare metal ‘real’ server
  - Ensure enough CPU, network, and IO bandwidth for the guests
  - Do not overload Dom0 with unnecessary packages, products and workload

- **Dom-U (Guests)**
  - Same traditional questions
    - How many CPUs do you need?
    - How much I/O and network bandwidth should be dedicate?
Oracle RAC on Oracle VM Best Practices (Cont’d)

- **Over-Committing**
  - Do not over-commit CPUs or memory unless workload allows for it

- **Configure Time synchronization**
  - Set guests to manage their own system clocks
    - set `xen.independent_wallclock=1` in `/etc/sysctl.conf`

- **Diagwait**
  - Set explicitly to ’13‘ for 10g and 11g Release 1
  - Not required for 11g Release 2
  - Not OVM specific
  - See Metalink Note 580296.1 for more information
In Dom0, configure two or more HBAs/NICs for redundancy
  - HBAs/NICs should be configured as multipath-ed devices in Dom0

Configure device persistency in Dom0
  - Use vendor multipathing or dm-udev in Dom0
  - Optionally configure physical block devices (in DomU) as ASMLIB disks
# xen config file example for RAC Guest Domain
name = "vmrac1"
memory

disk = ['phy:/dev/mapper/mpath3p1,xvda,w!',
'phy:/dev/mapper/mpath4p1,xvdb,w!',
'phy:/dev/mapper/mpath5p1,xvdc,w!'],

vif = ['mac=00:16:3E:00:00:08, bridge=xenbr0',
'mac=00:16:3E:10:A5:96, bridge=xenbr1'],

vfb = ['type=vnc,vncunused=1']

uuid = "3d6f1de4-626c-e02a-42a1-458c9c17e728"
bootloader="/usr/bin/pygrub"

vcpus=8

on_reboot = 'restart'
on_crash = 'restart'
RAC on OVM – Network Configuration

- Use dedicated network devices
  - For each host a minimum of two 1Gb Ethernet NICs
    - Public Network
    - Private Network
  - Extra Ethernet NICs required if network based storage connectivity (iSCSI or NFS) is used

- For complete network redundancy, 4-6 NICs recommended
- Bond ethx interfaces in Dom0 for HA
- These networks can be shared with multiple RAC VM guests, but network bandwidth and capacity needs to reviewed and correctly sized
RAC on OVM – Network Configuration

Guest 1

Public network

eth0

Private network

eth1

Domain-0

xenbr0

bond0

eth0

eth1

Public Network Switch 1

Public Network Switch 2

Private Network Switch 1

Private Network Switch 2

xenbr0

bond0

eth0

eth1
Oracle VM Templates
Rapid Application Deployment

Download from Oracle
- Pre-built, pre-configured VM
- Complete Clusterware, ASM, RAC installation
- Database 11g, Enterprise Manager dbControl

Save days or weeks in installation and configuration time

Oracle VM Server Pool

Oracle VM Servers

NAS, SAN, iSCSI

Customize & Save as Golden Images

Pre-built, pre-configured VM
- Complete Clusterware, ASM, RAC installation
- Database 11g, Enterprise Manager dbControl

Import to Oracle VM Manager

Start-Up in Oracle VM Pool
RAC on OVM – Best Practices Summary

• Each RAC node must be deployed on a separate physical server for production environments

• Be conservative if planning on over-committing CPUs or memory when running RAC on OVM.

• Configure disk multi-pathing and NIC bonding in Dom0

• To sustain required I/O throughput, ensure sufficient CPU resources are allocated to hypervisor (Dom0)

• Use the ready-made RAC OVM Templates!
See You in Oracle World 2011

• Oracle RAC OVM Templates download locations
  – 11.1.0.7 & 11.2.0.1 from e-delivery
    • 11.1: http://www.oracle.com/technetwork/database/clustering/overview/rac-092962.html
    • 11.2.0.1: http://www.oracle.com/technetwork/database/clustering/overview/rac-template-11grel2-166623.html
  – 11.2.0.2 from My Support: http://support.oracle.com/ Patch# 10113572
    – All described in Note:1185244.1:
      https://support.oracle.com/oip/faces/secure/km/DocumentDisplay.jspx?id=1185244.1

• Oracle RAC related sessions at Oracle World 2010

• Linux related sessions at Oracle World 2010
QUESTIONS & ANSWERS
SOFTWARE. HARDWARE. COMPLETE.
Oracle OpenWorld 2010

Roger Lopez
Systems Engineer
Oracle Database Solutions - Charter

• Our Charter – Deliver **robust, reliable, and scalable** solutions that solve customer problems while **removing complexity** and providing choice

• Complete **ownership** of the product design & development cycle which includes:
  › Integration
  › Validation
  › Bundling
  › Sustaining
Building Block Approach

- Software
  - Database
    - Enterprise Linux
    - Microsoft Windows 2003 SP2/2008/2008 R2
  - Database Applications
- Operating Systems
  - Operating Systems
  - Server
    - PowerEdge Servers (Rack / Tower / Blade)
  - Fabric
    - 10GbE/1GbE
    - Fibre Channel
    - SAS
    - SATA
  - Storage
    - iSCSI Storage
    - FC SAN

Global Marketing
Oracle Database Solutions - Deliverables

• **Create** Reference Architectures
• **Write** Best Practices
• **Deliver** quarterly solutions that include the following collaterals:
  › Deployment scripts
  › Deployment guides
  › Solution Deliverable List (SDL)
  › Oracle Solutions Advisor Tool found at [http://www.dell.com/oracle](http://www.dell.com/oracle)
How Dell uses Oracle VM Templates

• At Dell, we use Oracle VM Templates to do the following:
  › To build solutions around Oracle’s templates
  › To ease the installation process for development environments
  › To distribute templates for Dell IT development teams
  › To maintain a standard for deploying OS/Oracle software/Patching upgrades across Dell IT
To learn more about how Dell can help you drive an Efficient Enterprise visit:

- Dell’s onsite TSR for a free quote
- Michael Dell’s keynote on Wednesday at 8 a.m.
- One of Dell’s 20 conference sessions
- www.dell.com/oracle