ORACLE®

Power of the New Oracle RAC 11g Release 2 Oracle VM Templates

Saar Maoz & Philip Newlan
RACPack – RAC Development, Oracle

Updated: 23-NOV-2010

Agenda

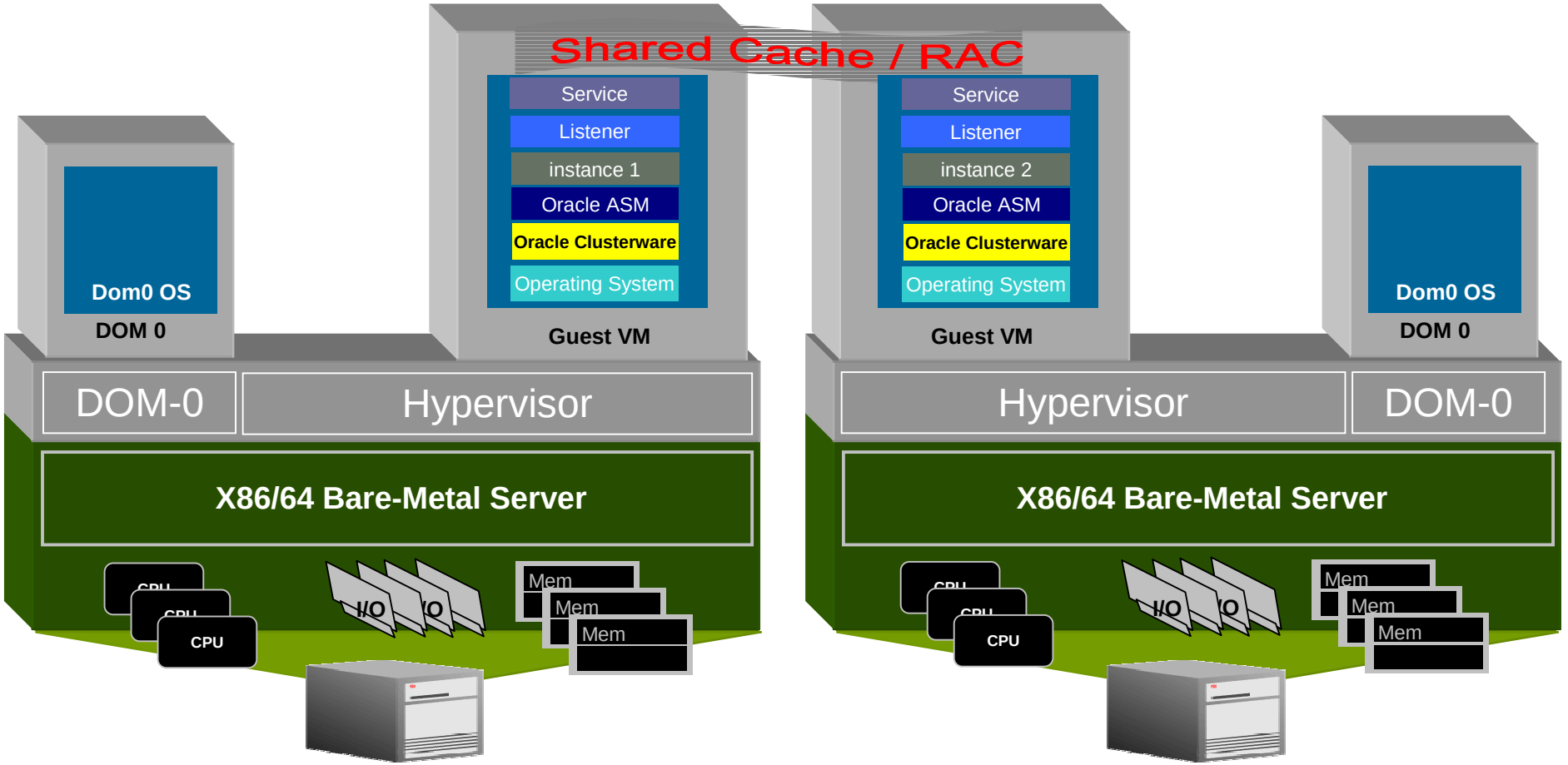
- Oracle RAC & Oracle VM Overview
- Oracle RAC – Oracle VM Templates
- Demos
  - Standard 2 node Cluster Build
  - Dom0 N-node Cluster Build
  - Add / Remove Nodes / Instances
  - Live Migration

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
  - Templates support both configurations, each has its own PDF to help with steps

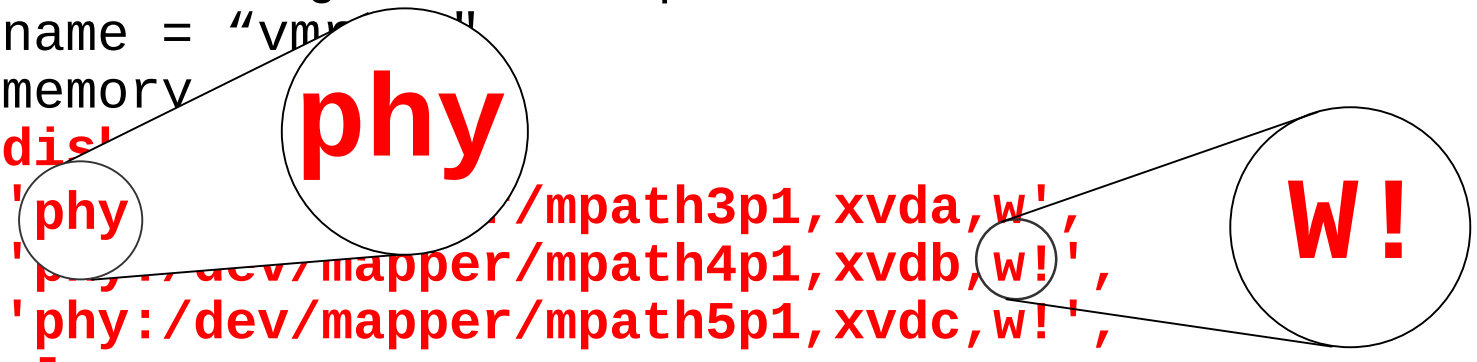
Oracle RAC on Oracle VM

Production Deployment

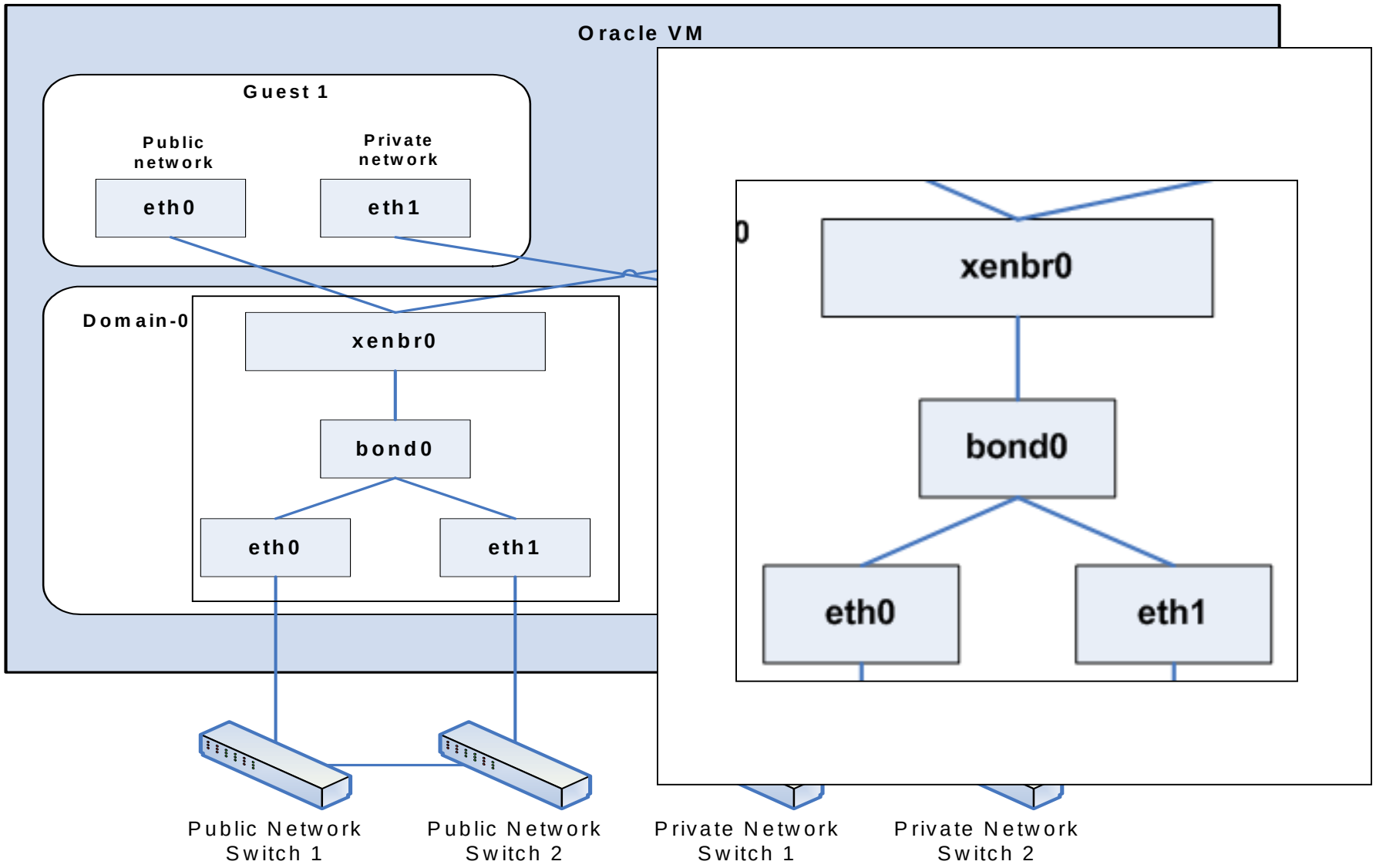


RAC on OVM – Disk Configuration (Production)

```
# xen config file example for RAC Guest Domain
  name = “vm
  memory
  dis
  'phy                    /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

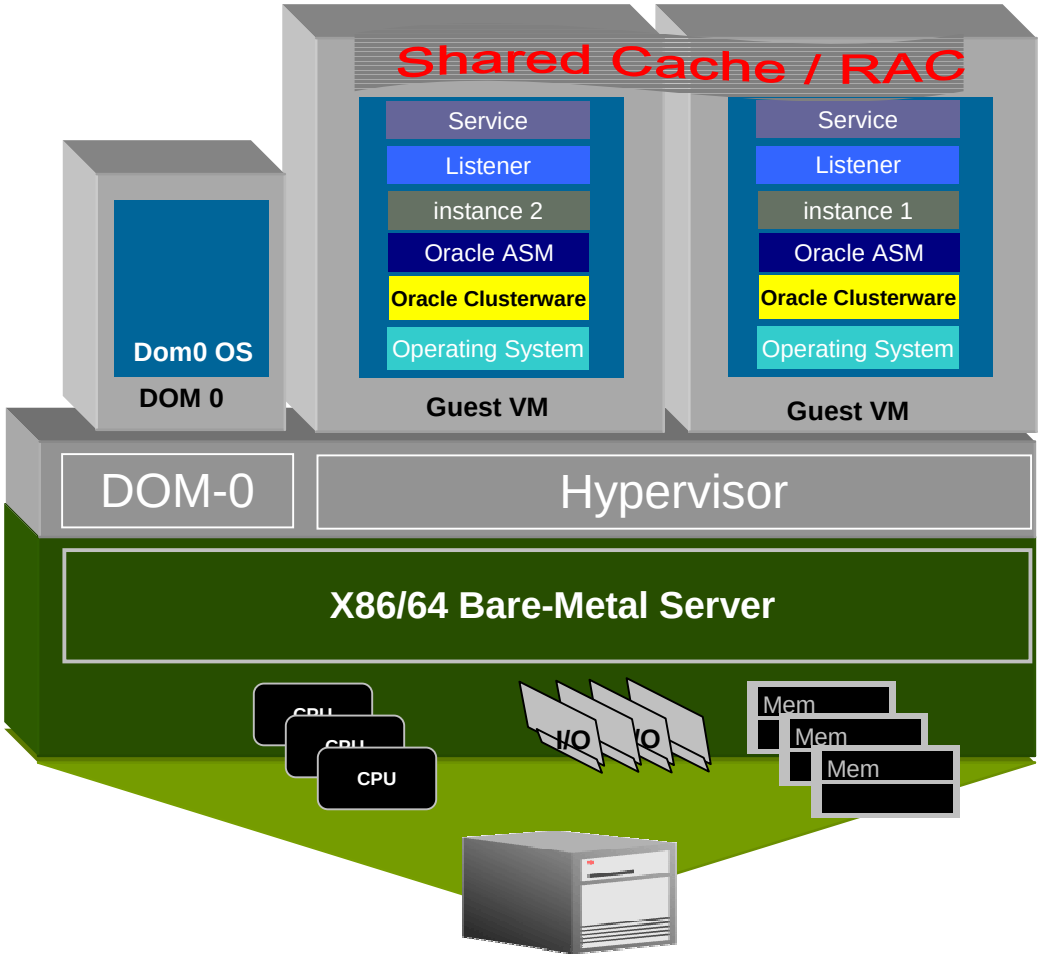


Oracle RAC on Oracle VM

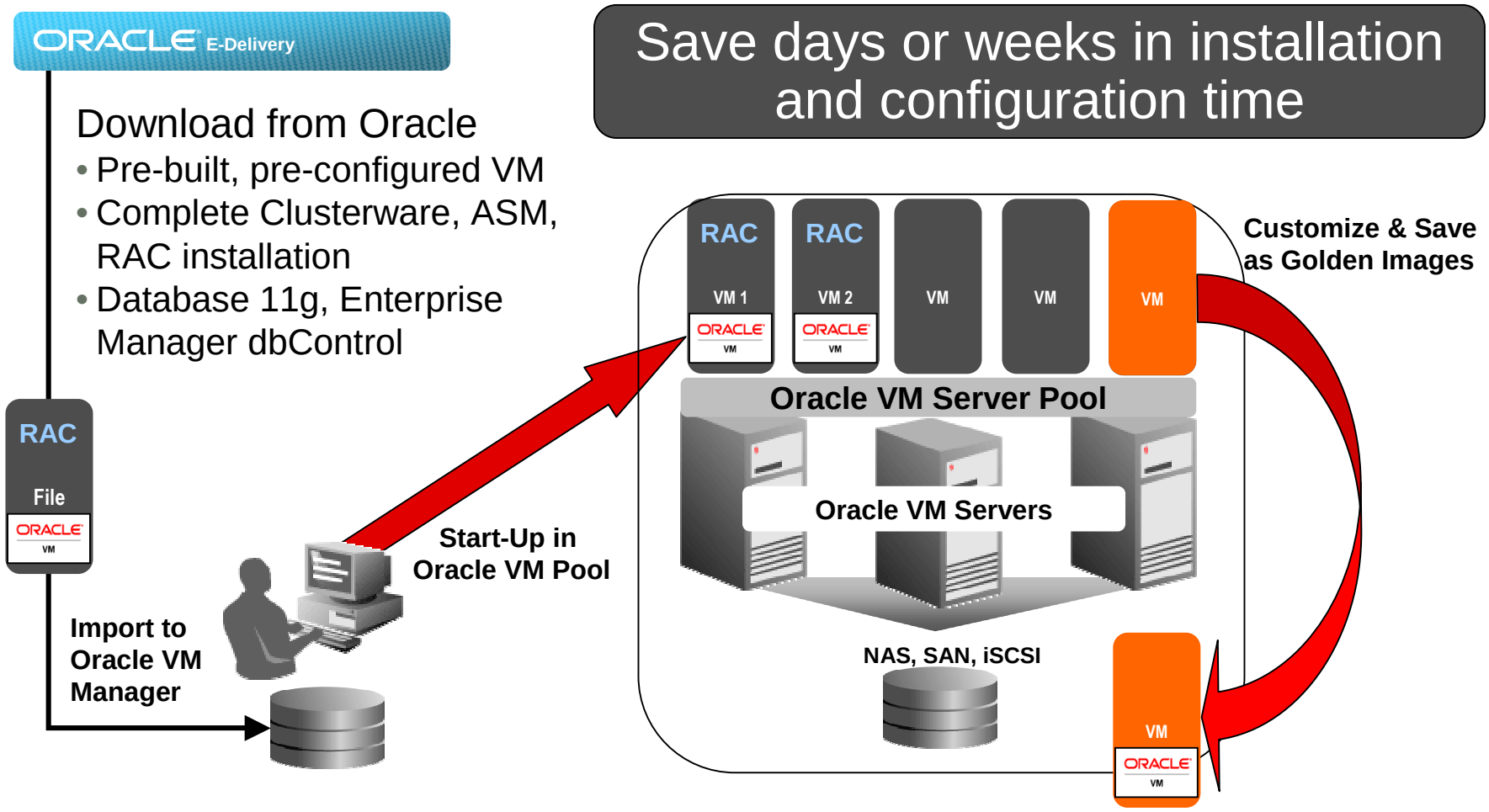
Non-Production deployment

2-node Test RAC Minimum Requirements

- Two or more cores
- 4GB of memory or more
- 30GB of disk or more



Oracle VM Templates
Rapid Application 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.jspx?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
- Follow the instructions in provided PDF files

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

```
NODE DETAILS              NODE 1                          NODE 2
============ -------------------------------   -------------------------------
Public Name:[test170                          ] [test171                        ]
Public IP   :[192.168.1.170   ]                   [192.168.1.171   ]
Private Name[test170-priv                     ] [test171-priv                   ]
Private IP  :[10.10.10.170    ]                   [10.10.10.171    ]
VIP Nodename[test170-vip                      ] [test171-vip                    ]
VIP IP      :[192.168.1.172   ]                   [192.168.1.173   ]
GLOBAL DETAILS
===============
Domain Name :[localdomain                      ] DNS Server IP :[               ]
Public Network Adapter Name :[eth0             ]
Public Subnet mask          :[255.255.255.0    ]
Default Gateway IP          :[192.168.1.1      ]
Private Network Adapter Name:[eth1             ]
Private Subnet mask         :[255.255.252.0    ]

CLUSTER DETAILS
===============
Enter the Cluster Name :[jun16-17071     ]
Enter the SCAN name for this cluster  :[scan-170171       ][192.168.1.174    ]


   DO YOU WANT TO CONFIGURE THIS CLUSTER YES/NO : >YES <
```

Building a Cluster

- After initial boot, to build a cluster, simply run:

```
/u01/racovm/buildcluster.sh
```

  - Automatically builds a cluster based on netconfig.ini & params.ini
  - Automatically writes a build log and progress log

- Internal testers reported:
  - 2 nodes, 18 minutes
  - 4 nodes, 23 minutes
  - End-to-end, including DBCA

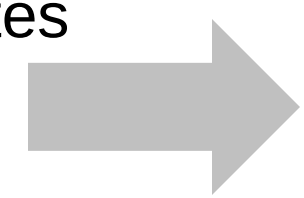
Configuration parameters

- `netconfig.ini` (Network information)
  - Node name/IP, (6 items per node)
  - NIC information
  - Cluster name
  - Mostly fixed data
- `params.ini` (Build related options)
  - Database name, user names, disks, etc.
  - More dynamic data

Building Clusters with more than 2 nodes

True Silent Install

- Customer supplies an initialisation file. (netconfig.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



netconfig.ini

```
# Node specific information
NODE1=test170
NODE1IP=192.168.1.170
NODE1PRIV=test170-priv
NODE1PRIVIP=10.10.10.170
NODE1VIP=test170-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
```

Building a Cluster from DOM-0

- Stamp netconfig.ini to shared disk:

  # **./netconfig.sh -W -c /dev/sdxyz**

  – If filesystem disk, use "**losetup –vf**" to loop mount the disk, then stamp the loop device.

- Boot VMs with node number hint (on Dom0):

  # **xm create -c vm.cfg extra="console=hvc0 template-reconfig-args='-R -n1 -b'"** (node 1)

  # **xm create -c vm.cfg extra="console=hvc0 template-reconfig-args='-R -n2'"** (node 2)

  -R read network config **-n1** node 1 **-b** build node

- Above will automatically configure the network on the new VMs and build a 2 node cluster

Running RACOVM directly

- List all steps:

  `/u01/racovm/racovm.sh -l`
- Running individual steps in sequence:

  `/u01/racovm/racovm.sh -S setsshroot,check``local`
- Most steps can be run global or local (add ‘local’ to step)
- Any failure of any step will stop execution
- Combination of common steps are also available as special steps, e.g. ‘buildcluster’ or command line flags, e.g. **`-c`**
- To cleanup run:

  `/u01/racovm/racovm.sh -S clean`

Running RACOVM directly (Cont'd)

- Each invocation prints this banner:

```
Invoking on test125 as root...
   Oracle RAC 11gR2 OneCommand (v1.1) for Oracle VM - (c) 2010 Oracle
   Cksum: [4028880687 234100 racovm.sh] at Mon Oct  4 14:14:14 EDT 2010
   Kernel: 2.6.18-164.0.0.0.1.el5xen (i686) [2 processor(s)] 1800 MB
   Step(s): setsshroot checklocal
```

- Timing for each operation:

```
2010-10-04 14:48:16:[buildcluster:Start:test235] Build 11gR2 RAC Cluster
2010-10-04 15:32:36:[buildcluster:Done :test235] Build 11gR2 RAC Cluster
2010-10-04 15:32:36:[buildcluster:Time :test235] Completed successfully in
2660 seconds (0h:44m:20s)
```

RACOVM Command Line Options

```
Usage: racovm.sh (v1.1)    RAC OVM main driver
       Oracle RAC 11gR2 OneCommand (v1.1) for Oracle VM - (c) 2010 Oracle
  -c : Builds a new RAC cluster including database                  [buildcluster]
  -g : Configures & starts Grid Infrastructure on all nodes calls:
       setsshroot,copykit,usrsgrps,printparams,setsshora,diskconfig,check,
       creategrid,cvupostcrs,clusterstate
  -r : Configures RAC Home by running clone.pl on all nodes         [racclone]
  -d : Run DBCA to create the RAC database                          [createdb]
  -p : Setup passwordless ssh for Oracle users between all nodes    [setsshora]
  -P : Setup passwordless ssh for root user between all nodes       [setsshroot]
  -S <stepname>,<stepname> ... : Run individual steps (comma separated list)
                                 stops execution on first error
  -N <node3>,<node4> .. : List of nodes to add or remove
      Required for 'addnodes' 'removenodes' 'addinstances' 'deleteinstances'
      Optional for 'createdb'
  -l : List all available steps
```

Adding or Removing Node(s) / Instance(s)

- Fully automated addition and removal of nodes or instances
- Simply run:

```
./racovm.sh -S addnodes -N node2,node3
```

Or:

```
./racovm.sh -S removenodes -N node2,node3
```

Or:

```
./racovm.sh -S addinstances -N node2,node3
```

Or:

```
./racovm.sh -S deleteinstances -N node2,node3
```

diskconfig.sh – Configures disks in VMs

- Verifies disks are not held on any node by
  - ASM, ASMLib, RAID device, PowerPath, Device Mapper, User Application, Filesystem, Swap Device
- Stamps and discovers disks on all nodes (verify sharedness)
- Auto-partition & align data to 1MB offset (default)
- Supports MSDOS or GPT partition table
- Merges needed udev rules to /etc/udev/…
  - Supports EL4, EL5 & SLES10, SLES11
- Supports but does not do the initial creation of
  - ASMLib
  - Multipath

netconfig.sh – Configures network in VMs

- Full validation on user input, NIC names, IP/subnet masks
- Checks for duplicate IPs on subnet (arping)
- Writes /etc/hosts and related ifcfg-*, resolv.conf, etc. files to fully configure network
- Allows stamping of `netconfig.ini` to shared storage; helps in N-node network configuration (from dom0 or inside guests)
- Supports and configures bonding (not needed inside guests)

doall.sh – Run command on all nodes

```
doall.sh [options] <command> | "<command1>;<command2>;..<commandX>"
-l <user>  : Login as user (default: current user)
-x         : Run commands in sequence (default: run in parallel)
-P         : Print commands as they are executed (default: noprint)
-p         : Setup passwordless SSH for <user> on all nodes listed in
             netconfig.ini
-s         : Skip confirmation for passwordless SSH setup
-u         : Skip unreachable nodes
-U         : Skip nodes that are unreachable or without passwordless
             SSH configured
-L first|last|skip    : Local node execution order
-N <node1>,<node2> .. : Nodes to run on (default: all nodes in
                        netconfig.ini)
```

Examples:

```
./doall.sh -L last  reboot

./doall.sh -ps /u01/app/11.2.0/grid/bin/diagcollection.sh
```

Useful Links

- 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 (requires support subscription)
  - All described in Note:1185244.1: https://support.oracle.com/oip/faces/secure/km/DocumentDisplay.jspx?id=1185244.1


Q&A
QUESTIONS
ANSWERS

Hardware and Software

ORACLE

Engineered to Work Together