Oracle RAC 10g
over
Veritas SFRAC

by
Ashok Kapur
Hawkeye Technology, Inc.
http://www.hawkeyetechnology.com

Agenda

- **RAC Introduction**
- **Architecture**
- **Installation & Verification Steps**
- **Configuration**
- **Best Practices**
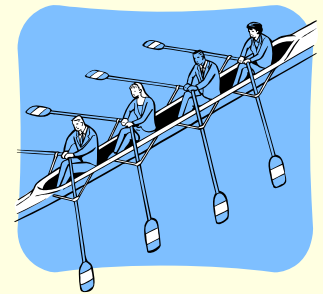- **Questions?**

RAC Introduction

Real Application Cluster

- RAC: Real Application Cluster
  - Multiple servers (nodes) act as a single "clustered" server.
  - Multiple database instances (one on each server) accessing the same database over shared storage.
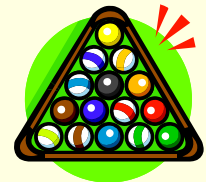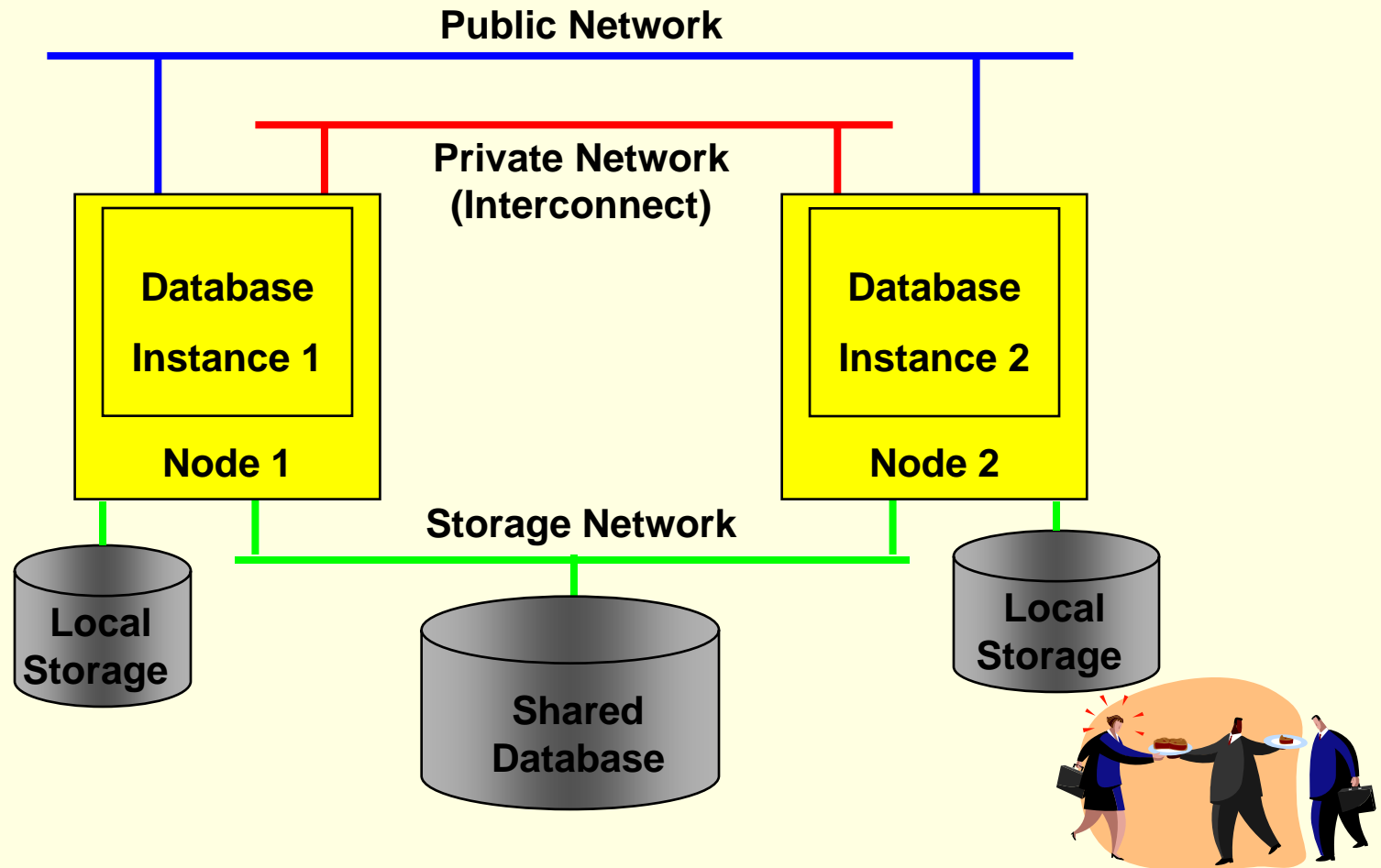
Pros/Cons

- Pros
  - High Availability
  - Scalability
  - Load Balancing/Sharing
- Cons
  - Increased Complexity
  - Increased Cost

2 Node RAC

Architecture

Oracle RAC over Veritas SFRAC

Architecture

App Server 1 · App Server 2

Public Network

Node 1 VIP · Node 2 VIP

DB Server 1 · DB Server 2

Private Interconect

Shared Storage:
Database
I/O Fencing Disks

RAC Communication

Oracle RAC over Veritas SFRAC
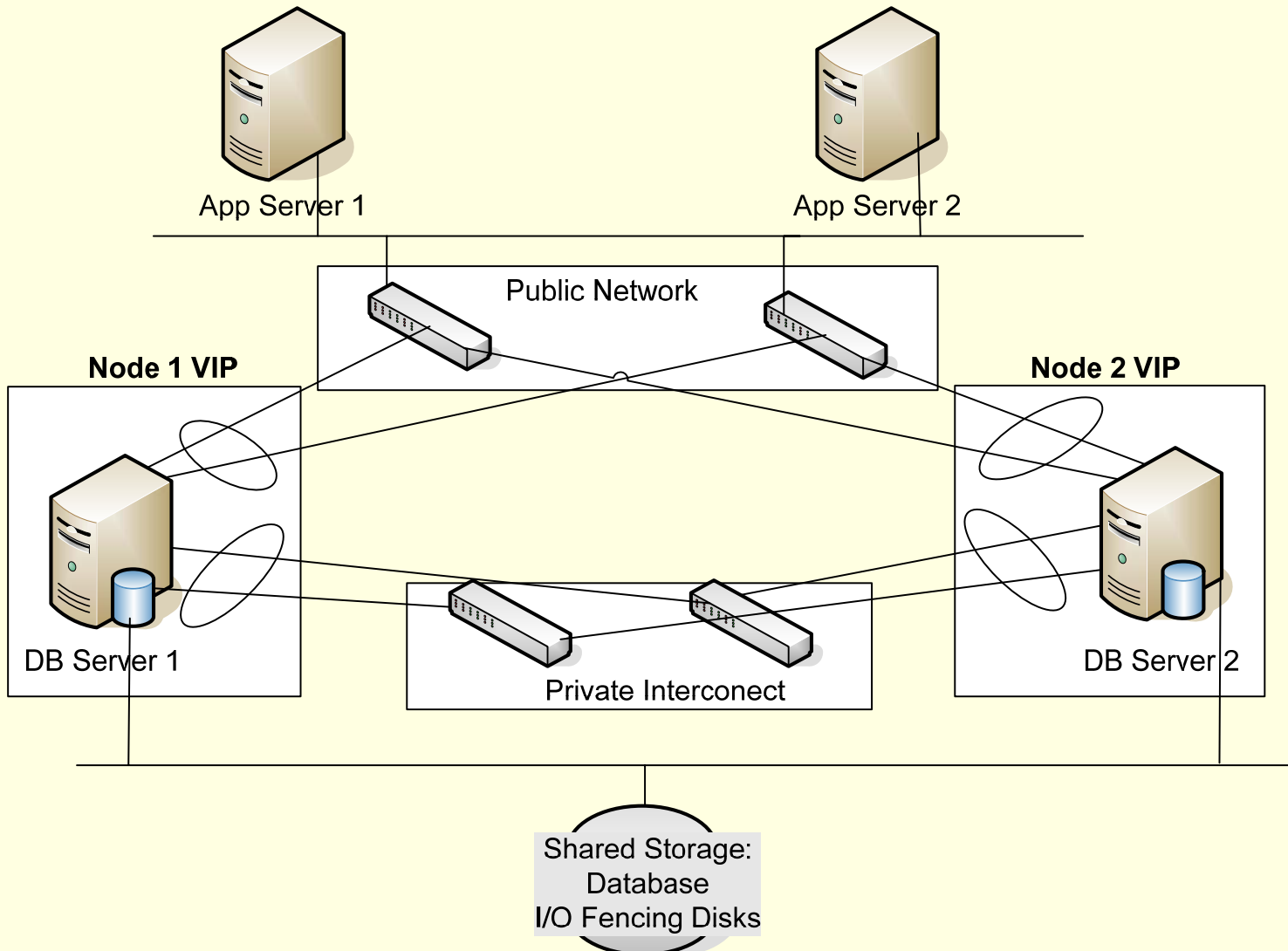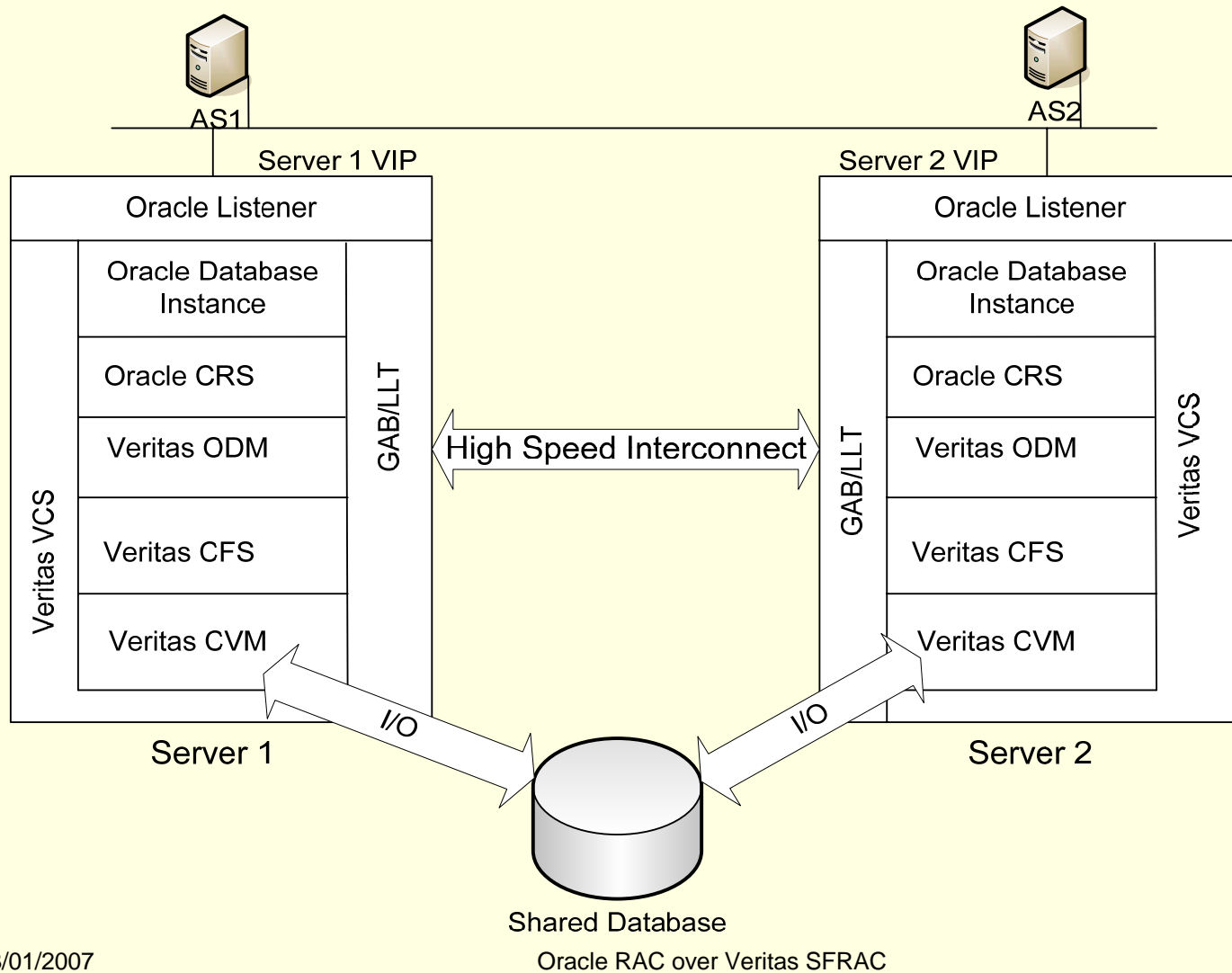
RAC Communication

- Low Latency Transport (LLT)
- Group Membership Srvc/Atomic Broadcast (GAB)
- Cluster Volume Manager (CVM)
- Cluster File System (CFS)
- Oracle Disk Manager (ODM)
- Veritas Cluster Server (VCS)
- Cluster Ready Service (CRS)
- Oracle Database Instance
- Oracle Listener

LLT

- LLT provides fast kernel-to-kernel communications

- Proprietary Veritas communication protocol

- Runs over DLPI (Data Link Protocol Interface) layer.

- LLT load-balances communications over Cluster interconnects.

- Responsible for sending/receiving heartbeat over the interconnects.

GAB

- Manages node membership in the cluster.
- Uses LLT heartbeat to determine cluster membership.
- Provides guaranteed delivery of all messages over the cluster interconnect.

CVM

- Extension of VxVM (Veritas volume manager) for Clusters.

- Uses Master/Slave concept.

- Changes to master node configuration are propagated to all slave nodes.

- CVM does not impose any write locks among the members. Data integrity is the responsibility of the upper application layer (CFS).

- CVM imposes all-or-none rule whereby each node must connect to all disks of a given diskgroup.

CFS

- CFS enables a file system to be mounted on multiple servers at the same time.
- All nodes have a consistent view of data

ODM

- API that Oracle uses for all I/O.
- ODM bypasses GLM (Global Lock Manager) and data buffers in CFS.
- ODM is reported to provide RAW device access speeds.

VCS

- Monitors and manages processes under its control.

- Responsible for startup/shutdown and monitoring of processes.

- If a process dies, it restarts it.

CRS

- CRS (Cluster Ready Services) is a new feature for 10g Real Application Clusters

- Provides a standard cluster interface on all platforms and performs new high availability operations not available in previous versions.

- Can be installed by itself or on top of vendor Cluterware.

- Can performs all Cluster management (global lock management, split brain, cache fusion,…) functions by itself.

- Need Veritas for Cluster File System

CRS Processes

- CRSD: Engine for HA operation
  - Manages (start/stop/respawn) application resources
  - Maintains configuration profiles in the OCR (Oracle Configuration Repository)
  - Stores current known state in the OCR
  - Runs as root
  - Is restarted automatically on failure
- OCSSD:
  - OCSSD is part of RAC and Single Instance with ASM
  - Provides access to node membership, group services, basic cluster locking
  - Integrates with vendor clusterware, when present
  - Can also runs without integration to vendor clusterware
  - Runs as Oracle
  - Failure exit causes machine reboot
  - Prevents data corruption in event of a split brain.
- EVMD:
  - Generates events when things happen
  - Spawns/respawns evmlogger(s)
  - Scans callout directory and invokes callouts
  - Runs as Oracle
  - It is respawned automatically on failure
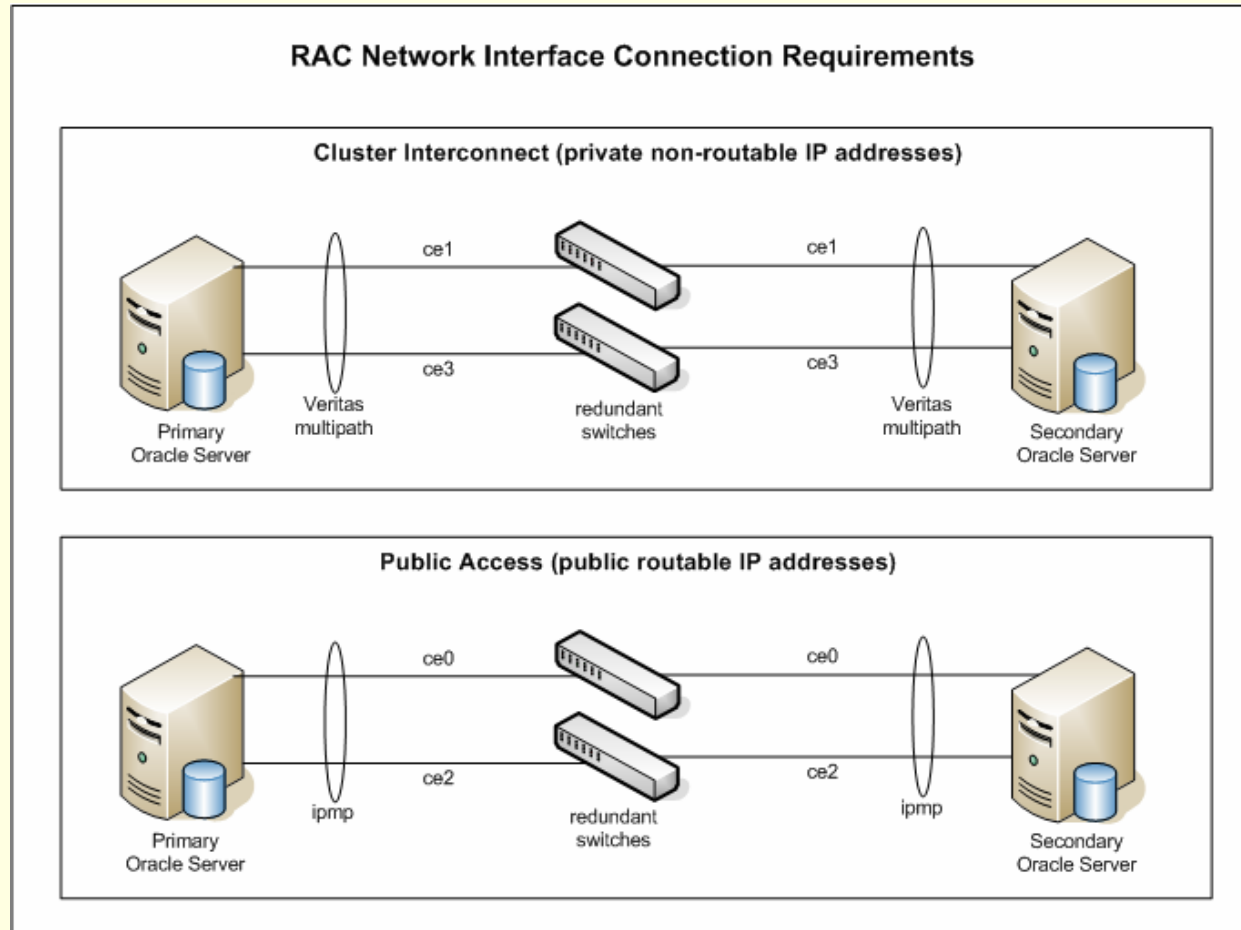
Installation

Installation Steps

- Pre-requisites
- Install SFRAC 4.1
- Configure IO Fencing
- Install SFRAC 4.1 MP1
- Post SFRAC Install Steps
- Pre-CRS Install Steps
- Install CRS
- Install Oracle 10g
- Create Database

Pre-Requisites

- Ensure eeprom local-mac-address?=true on both nodes
- Setup export PATH=$PATH:/sbin:/opt/VRTSvcs for root.
- Change hostnames for all servers to ALL lowercase.
- Set interface speed to 1000/full for interconnect interfaces and 100/full for public interfaces.
- Setup ssh and scp without a password.
- Acquire INTERNAL IP addresses: one per node.
- Three additional PUBLIC IP addresses are needed:
    - 2 cluster virtual IPs (one for each node) registered in DNS
    - 1 Veritas GUI Management virtual IP, registered in DNS
- Setup IPMP between two public IP addresses (two public NICS per node setup as IPMP)
- Plumb both public interfaces on each node.
- The two cluster-interconnects need to be on different switches and on different VLANs.
- Follow "Chapter 7 - Pre-Installation Tasks for RAC on Solaris" of the *Oracle Real Application Clusters Installation and Configuration Guide*.
- Ensure /etc/system is identical for both servers
- LOCAL ORACLE_HOME and CRS_HOME
- Follow pages 24-27 of the *VERITAS Storage Foundation 4.1 for Oracle RAC: Installation and Configuration Guide*.
- Setup three disks in SAN as coordination disks. Since there is no data stored on these disks, these can be configured as the smallest possible LUNs on the disk array.
- Ensure all SAN disks to be used for shared disk storage are setup as SCSI-3 disks and support SCSI-3 persistent reservations. This includes I/O fencing disks.

RAC Network Connections

RAC Network Interface Connection Requirements

Cluster Interconnect (private non-routable IP addresses)

ce1 — ce1 — Veritas multipath — redundant switches — Veritas multipath
ce3 — ce3

Primary Oracle Server — Secondary Oracle Server

Public Access (public routable IP addresses)

ce0 — ce0
ce2 — ce2
ipmp — redundant switches — ipmp

Primary Oracle Server — Secondary Oracle Server

Install SFRAC

- Follow "VERITAS Storage Foundation 4.1 for Oracle RAC: Installation and Configuration Guide" for install instructions with following variations/changes:
    - Copy installation CD3 onto disk and start install from disk. (pg 30-31)
    - Start the installer using the switch **./installer –usessh** in order to use ssh during the install (pg 31)
    - Select I to Install
    - Select 7 to install "VERTIAS Storage Foundation for Oracle RAC"
    - Press [Return] to continue
    - Select "y" to configure SFRAC in step 11, pg 35.
    - Select "n" for step 3, pg 37.
    - Select "y" to configure Cluster Manager in step 4, pg 38. Enter a VIP to be used by Veritas Web Console.

Verify SFRAC

- # gabconfig –a

GAB Port Memberships

================================

Port a gen   3a6502 membership 01

Port d gen   3a6501 membership 01

Port h gen   3a6506 membership 01

Port o gen   3a6504 membership 01

GAB Port Cross-reference

GAB Port	Function
a	GAB driver
b	I/O fencing (designed to guarantee data integrity)
d	ODM (Oracle Disk Manager)
f	CFS (Cluster File System)
h	VCS (VERITAS Cluster Server: high availability daemon)
o	VCSMM driver (kernel module needed for Oracle and VCS interface)
q	QuickLog daemon
v	CVM (Cluster Volume Manager)
w	vxconfigd (module for cvm)

Configure IO Fencing

- Stop VCS.

  - **# hastop –all**

- Follow instructions on pages 45-60 of the "VERITAS Storage Foundation 4.1 for Oracle RAC: Installation and Configuration Guide", except:

  - In step 2 on pg 47, use **vxfentsthdw -s** to use ssh instead of rsh

  - In step 1 on pg 50 use **vxfentsthdw -s -c** to use ssh instead of rsh

  - Do not remove ssh setup as instructed on pg 53.

  - Reboot the servers

Verify IO Fencing

- # /sbin/gabconfig –a

GAB Port Memberships

```
===================================
Port a gen   1601206 membership 01
Port b gen   160120a membership 01
Port d gen   1601208 membership 01
Port f gen   1601218 membership 01
Port h gen   160120e membership 01
Port o gen   160120c membership 01
Port q gen   1601216 membership 01
Port v gen   1601212 membership 01
Port w gen   1601214 membership 01
```

Install SFRAC 4.1 MP1

- Copy MP1 files on disk.
- Verify current version of SFRAC packages on page 10.
- Stop llt and gab
- NOTE: llt drivers and gab drivers do not unload properly. Follow the following directions to stop these <u>prior to starting the install</u>:
  - On each node, logged in as root issue: *mv S70llt s70llt.*
  - Reboot each server.
- Verify llt is not running by issuing:
- # gabconfig -a
  GAB Port Memberships
  #
- Install MP1 as directed on page 11
  - Use # ./install_vp –usessh command to start the installation
- Reboot the server after the installation of MP1.

Verify MP1 Installation

■ After the system is rebooted, run the following command on both the nodes to verify that both nodes have memberships to ports a, b, d, f, h, o, q, v and w:

/sbin/gabconfig –a

```
GAB Port Memberships
===================================
Port a gen  1601206 membership 01
Port b gen  160120a membership 01
Port d gen  1601208 membership 01
Port f gen  1601218 membership 01
Port h gen  160120e membership 01
Port o gen  160120c membership 01
Port q gen  1601216 membership 01
Port v gen  1601212 membership 01
Port w gen  1601214 membership 01
```

■ Verify the CVM group is online:

hagrp –state cvm

Group	Attribute	System	Value
cvm	State	udbm03320-01-01-01-008 |ONLINE|
cvm	State	udbm03320-01-01-01-006 |ONLINE|

■ Verify VEA service is running:

/etc/init.d/isisd status
Current state of server : RUNNING

Post SFRAC

- Create Shared Volumes
- Create Oracle account
- Create local directories for
  - ORALCLE_HOME
  - CRS_HOME
  - ORACLE_BASE/admin

Pre-CRS Installation

- Login as user "oracle"
- Follow instructions on pages 83-95 of the "VERITAS Storage Foundation 4.1 for Oracle RAC: Installation and Configuration Guide". Note the following exceptions:
    - On page 88 under "Copying SFRAC Libraries…" add:
    - # cp /opt/VRTSvcs/rac/lib/libskgxp10_64.so /opt/ORCLcluster/lib/libskgxp10.so
- Setup environment variables:
    - ORACLE_BASE=/opt/oracrs
    - ORACLE_HOME=$ORACLE_BASE/product/10.1.0
    - LD_LIBRARY_PATH64=$ORACLE_HOME/lib
    - LD_LIBRARY_PATH=$ORACLE_HOME/lib32
    - PATH=$PATH:$ORACLE_HOME/bin
    - CLASS_PATH=$CLASSPATH:$ORACLE_HOME/JRE:$ORACLE_HOME/jlib:$ORACLE_HOME/rdbms/jlib:$ORACLE_HOME/network/jlib
- Copy SFRAC libraries to /opt/ORCLcluster/lib

    # cp /opt/VRTSvcs/rac/lib/libskgxn2_64.so /opt/ORCLcluster/lib/libskgxn2.so
    # cp /opt/VRTSvcs/rac/lib/libskgxp10_64.so /opt/ORCLcluster/lib/libskgxp10.so
- Run the "cluvfy" utility as root to ensure all pre-requisites are met, and resolve any issues before proceeding:

    # ./cvufy stage –pre crsinst –n <node1>,<node2> -verbose | tee /tmp/cvuverify_preinst.log

** NOTE: Use CVU (Oracle's Cluster Verification Utility) throughout the CRS Installation!!

CRS Installation

- Install CRS per Oracle's install instructions.
- Nodes and network interfaces should be listed during the install.
- **WARNING: Prior to running the root.sh script**, do the following:
    - Patch the init.cssd script as follows:
        - Log in as the root user, from a separate window
        - Change to the directory where the patch is to be copied, and make a backup of the original version init.cssd file:

            # cd $ORA_CRS_HOME/css/admin

            # cp init.cssd init.cssd.original

            # chmod 755 init.cssd

            # cp /opt/VRTSvcs/rac/patch/init.cssd.patch .
        - Run the following command to install the patch:

            # patch init.cssd < init.cssd.patch
        - Verify patch installation by checking init.cssd to see that it contains the following:

            DFL_CLSINFO=/opt/VRTSvcs/ops/bin/clsinfo

Verify CRS Install

- Run the following utility, which should exit without errors:
$ORA_CRS_HOME/bin/crs_stat

- Run the following command, which should list the nodes in the cluster:
$ORA_CRS_HOME/bin/olsnodes –n

- Run the following command, which should return "CSS daemon appears healthy"
$ORA_CRS_HOME/bin/crsctl check css

- Run the following command and review output for any errors or anomalies reported:
$ORA_CRS_HOME/bin/cluvfy stage –post crsinst –n all –verbose | tee /tmp/cluvfy_postcrs.log

Install Oracle 10g

- Pre-Install: Run CVU
    - ./cvufy stage –pre dbinst –n <node1>,<node2> -verbose | tee /tmp/cvuverify_predbinst.log
- Install Oracle 10g per Install Instructions
- Post-Install:
    - Login as user "oracle" on one node
    - Link SFRAC libraries into ORACLE_HOME:
      $ cd $ORACLE_HOME/lib
      $ cp libskgxn2.so libskgxn2.so.orig
      $ cp libskgxp10.so libskgxp10.so.orig
      $ mv libodm10.so libodm10.so.orig
      $ cp /opt/VRTSvcs/rac/lib/libskgxn2_64.so libskgxn2.so
      $ ln –s libskgxn2.so libskgxn10.so
      $ cp /opt/VRTSvcs/rac/lib/libskgxp10_64.so libskgxp10.so
      $ ln –s /usr/lib/sparcv9/libodm.so libodm10.so
    - Repeat on all other nodes

Create and Register Database

- Create or migrate database
- Verify cluster membership via:
    - Select * from v$active_instances;
- Register database and listeners in SRVCTL
- Create/Update server and client tnsnames.ora

Configuration

Oracle RAC over Veritas SFRAC

Server TNSNAMES

```
SERVDB1 =
 (DESCRIPTION =
  (ADDRESS = (PROTOCOL = TCP)(HOST = node0-vip)(PORT = 1522))
  (CONNECT_DATA = (SERVER = DEDICATED) (SERVICE_NAME = SERVDB1)  )
 )

INST1 =
 (DESCRIPTION =
  (ADDRESS = (PROTOCOL = TCP)(HOST = node0-vip)(PORT = 1522))
  (CONNECT_DATA =   (SERVER = DEDICATED)   (SID = INST1)  )
 )

INST2 =
 (DESCRIPTION =
  (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1522))
  (CONNECT_DATA =  (SERVER = DEDICATED)   (SID = INST2)  )
 )
```

Initial INIT.ORA

…
*.cluster_database=true
*.cluster_database_instances=2
*.compatible='10.1.0.4'
*.undo_management='AUTO'
*.undo_tablespace='UNDOTBS1'
INST1.undo_tablespace='UNDOTBS1'
INST1.local_listener='(ADDRESS=(PROTOCOL=TCP)(HOST=node0-vip)(PORT=1522))'
INST1.remote_listener='(ADDRESS=(PROTOCOL=TCP)(HOST=node1-vip)(PORT=1522))'
INST1.thread=1
INST1.instance_number=1
INST2.undo_tablespace='UNDOTBS2'
INST2.local_listener='(ADDRESS=(PROTOCOL=TCP)(HOST=node1-vip)(PORT=1522))'
INST2.remote_listener='(ADDRESS=(PROTOCOL=TCP)(HOST=node0-vip)(PORT=1522))'
INST2.thread=2
INST2.instance_number=2

Client TNSNAMES

```
SERVDB1 =
 (DESCRIPTION =
  (ADDRESS_LIST=
    (FAILOVER=on)
    (LOAD_BALALCE=off)
    (ADDRESS = (PROTOCOL = TCP)(HOST = node0-vip)(PORT = 1522))
    (ADDRESS = (PROTOCOL = TCP)(HOST = node1-vip)(PORT = 1522))
   )
  (CONNECT_DATA =
    (SERVER = DEDICATED)
    (SERVICE_NAME = SERVDB1)
   )
 )
```

Best Practices

Best Practices

- Ensure ALL hostnames are lower case
- DO NOT cross-connect cluster interconnects
- Do not put Oracle services in Veritas VCS
- ORACLE_HOME, CRS_HOME on local disks with identical paths on all nodes
- Use CVU for CRS installation
- DO NOT change default LISTENER names given via NETCA
- Always patch CRS before patching ORACLE_HOME
- After every patch to ORACLE_HOME and before running root.sh, copy the Veritas libraries.
- Specify REMOTE_LISTENER in spfile

References

- VERITAS Software Foundation 4.1 for Oracle RAC, Installation and Configuration Guide, Solaris
- VERTIAS Storage Foundation 4.1 for Oracle RAC, Release Notes, Solaris, Maintenance Pack 1
- VERTIAS Storage Foundation 4.1 for Oracle RAC, Release Notes, Solaris
- Integrating Veritas Storage Foundation for Oracle RAC with Oracle 10g CRS, Solaris and HP-UX platforms
- Oracle Real Application Cluster Administration Guide
- Oracle RAC Installation and Configuration Guide

References (contd.)

- Oracle Metalink Notes:
  - **Note:296878.1**: **Oracle 10g VIP (Virtual IP) changes in Oracle 10g 10.1.0.4**
  - **Note:259301.1**:**CRS and 10g Real Application Clusters**
  - **Note:283107.1**: **Configuring Solaris IP Multipathing (IPMP) for the Oracle 10g VIP**
  - **Note:239998.1**: **10g RAC: How to Clean Up After a Failed CRS Install**

Questions/Comments?

CONTACT INFORMATION

Ashok Kapur

Hawkeye Technology, Inc.

afkapur@hawkeyetechnology.com