Oracle RAC Interview Questions

1. **What is Oracle RAC and its advantages?**

   RAC or Real Application Cluster is an option provided by Oracle Corporation that enables the oracle to be installed across multiple servers that access a shared storage.

   Oracle RAC uses Oracle Clusterware for the servers to group, so they operate as a single system.

   Oracle Real Application Clusters (RAC) provides high availability (HA) and scalability to the Oracle Database without requiring any application changes.

   Advantages:-

   - **Lower overall cost of ownership** – Oracle keeps the cost of ownership by reducing the administrative overhead.
   - **Expanded Scalability** – Oracle enables you to improve performance and add capacity by adding nodes. Some platforms supports node addition while cluster is running.
   - **High availability** – Multiple oracle instances provides uninterrupted services even during failures.
   - **Transparency** – It offers the same functional equivalency as single node oracle database.
   - **Rolling patching -** Some patches to the Oracle software allow for application one node at a time without a total database outage.

2. **What are the architectural difference between RAC and non-RAC (Standalone) databases?**

Component	Standalone Database	RAC Database
SGA	Dedicated to the instance	Each instance has its own SGA
Background process	Dedicated to the instance	Each instance has its own processes
Datafiles	Dedicated to the instance	Accessed by all instances (Shared Storage)
Control files	Dedicated to the instance	Accessed by all instances (Shared Storage)
Redo log files	Dedicated to the instance	Only one instance can write at a time but others can read data during recovery or archiving process.
Archive log files	Dedicated to the instance	Each instance creates archive logs
Flash recovery log	Dedicated to the instance	Accessed by all instances (Shared Storage)
Alert log and trace	Dedicated to the instance	Each instance has alert log and trace files
Undo Tablespace	Dedicated to the instance	Each instance has own UNDO tablespace

3. **What are the components of Oracle RAC?**
    - Clusterware (CRS)
    - Shared Storage (ASM)
    - Node/Database (RDBMS)
    - Interconnect

4. **Which are the major tools used in RAC?**
    - Cluster verification utility (cluvfy)
    - Oracle Enterprise Manager
    - SQL*plus
    - SRVCTL (Server Control Utility)
    - Database Configuration Assistant
    - Fleet patching and provisioning

5. **Which are the RAC specific parameters?**
    - undo_tablespaces
    - cluster_database
    - cluster_interconnects
    - remote_listener
    - thread
    - cluster_database_instances

6. **Why RAC database has separate redo log and UNDO tablespace for each instance?**

    Each instance maintains own memory and background, hence LGWR process has to write the changes from redo log buffer to online redo log files.

    As data blocks are being updated by sessions connected to each instance, oracle has to maintain instance specific UNDO tablespace to ensure read consistency.

7. **What are the type of IP addresses used in Oracle RAC and its purposes?**

    Public IP: Used for normal network communications to the node

    Private IP: Used as the cluster interconnect (Node to node communication)

    Virtual IP: Used for failover between nodes in the cluster.

    Public/Virtual/SCAN IPs should be in the same subnet as it is public facing

8. **Why do we need VIPs in Oracle RAC?**

- **High- availability -** Each node in an Oracle RAC cluster is assigned a VIP in addition to its regular IP address. In case of node failure, the VIP can be moved to a surviving node, ensuring continuous availability of the database services.
- **Fast Failover -** By avoiding lengthy TCP timeouts, VIPs enable quicker client reconnection during failovers.
- **Connection redirect -** VIPs facilitate Transparent Application Failover (TAF) and Fast Connection Failover (FCF), enabling automatic client reconnections to another node without manual intervention.

9. **What is ONS (Oracle Notification Service)?**

ONS is a daemon process that communicates with other ONS daemons on other nodes which inform each other of the current state of the database components on the database server.

10. **What is Transparent Application Failover (TAF)?**

TAF allows application to reconnect to the available service when the initial connection fails. Normally when the connection is failed, active transactions will be rolled back. However, using this method, SELECT statement can be carried over to execute from the other node.

Two methods are there:

a. BASIC – Reconnection will be happened when the initial session is failed.
b. PRECONNECT – Along with the initial connection, another shadow session will be established.
c. SESSION FAILOVER: Recreates lost connections and sessions.
d. SELECTFAILOVER: Replays queries that were in progress.

TAF can be achieved by two methods.

a. server-side configuration

srvctl add service -db RACDB -service APSVC -failovermethod BASIC -failovertype SELECT -failoverretry 10 -failoverdelay 30 -serverpool sp1

SELECT inst_id, username, service_name, failover_type, failover_method FROM gv$session WHERE username='AP';

b. client-side configuration

```
AP =
  (DESCRIPTION =(FAILOVER=ON)(LOAD_BALANCE=ON)
   (ADDRESS=(PROTOCOL=TCP)(HOST=N1VIP)(PORT=1521))
    (ADDRESS=(PROTOCOL=TCP)(HOST=N2VIP)(PORT=1521))
      (CONNECT_DATA =
        (SERVICE_NAME = AP)
```

(FAILOVER_MODE= (TYPE=select)

(METHOD=basic)

(RETRIES=20)

(DELAY=15))))

11. How do we ensure that the time remains synchronized on all nodes of a cluster?

It can be achieved through two ways

a. NTP (Network Time Protocol)
b. CTSS (Cluster Time Synchronization Service)

```
cd $GRID_HOME/bin
./cluvfy comp clocksync -n all

crsctl check ctss
```

12. How crash recovery and instance recovery can be explained?

Crash recovery – It takes place in single node database on startup when instance crashes.

Instance recovery – It takes place in RAC databases. When an instance is crashed, the same recovery for the instance is performed by surviving node.

13. Explain about local_listener and remote_listener parameters in RAC?

Local_listener in Oracle RAC is configured with Node VIP.

Remote_listener is configured with SCAN.

14. What is SCAN in oracle RAC?

Single Client Access Name (**SCAN**) is a feature used in Oracle RAC environments that provides a single name for clients to access any Oracle Database running in a cluster. The main benefit is that the client's connect information does not need to change if you add or remove nodes or databases in the cluster.

15. How many SCAN Listeners are needed for a 5 node RAC?

3 SCAN listeners would be sufficient. It is not mandatory that SCAN listeners should run in every nodes.

16. How a connection is established in Oracle RAC?

1. The client uses a Single Client Access Name (SCAN) to connect to one of the RAC database's listeners. SCAN acts like a load balancer, allowing clients to connect to the listener that's least loaded.
2. The SCAN listener receives the connection request and identifies the least loaded instance in the cluster that can provide the requested service.
3. The SCAN listener then redirects the connection request to the local listener on the node where the least loaded instance is running. The SCAN listener sends the local listener's address and port details to the client.
4. The client connects to the local listener, which creates a dedicated server process. The client then connects directly to the dedicated server process and accesses the database instance.

17. Which are the ways to configure SCAN?

- DNS (Domain Name Server) – Configure SCAN name in DNS to resolve IPs and balance the load using round robin algorithm.
- GNS (Oracle Grid Naming Service) – GNS will acquire required IP from DHCP and provides name resolution.
- Hosts file – This option can be used when no DNS server set up is in place.

18. How to check the status of SCAN and SCAN listener in oracle?

srvctl start/stop/status scan

srvctl start/stop/status scan_listener –i 3

srvctl relocate scan_listener –i 3 –n node_name -> to relocate a SCAN listener to another node.

srvctl modify scan_listener -u -p TCP:1521 -> to modify the port of a SCAN listener

19. How SCAN knows which node has least load?

Load balance Advisory provides load information to scan.

20. Can we have multiple SCAN names in a RAC set up?

From 12c onwards, we can have multiple scan with different subnets. As part of installation only scan will be configured. Post installation we need to configure another SCAN with different subnet.

21. How to find the cluster name, servers in the cluster and grid version in RAC?

olsnodes –c -> Cluster name

olsnodes -> Servers in the cluster

crsctl query crs softwareversion node_name -> grid version

22. Which files are responsible to store clusterware configurations?

Shared files: OCR and Voting disk

Local Files: GPnP Profile and OLR

23. What is votedisk and how to find its location?

Oracle clusterware uses voting disk to maintain the node membership information within the cluster.

The Cluster Synchronization Services daemon (CSSd) performs all the operations for the voting disk in Clusterware.

It is suggestable to place the voting disk in ASM diskgroup or in shared storage.

Always keep multiple copies of voting disk and we can have up to 32 copies of voting disk.

From 11g onwards, no manual backup of voting disk is needed.

Voting disk stores two types of data

1. Static data – information about nodes in the cluster
2. Dynamic data – Disk heartbeat logging

Following commands can be used.

To check the location of vote disk: crsctl query css votedisk

24. Why should we have odd number of votedisk?

A node must be able to access more than half of the voting disks at any time. When a node is not able to access more than half of the voting disk will be evicted from the cluster.

25. How to replace the corrupted voting disk?

a. Corrupt the voting disk
b. Shutdown the database
c. Stop the CRS service on all nodes
d. Start the CRS service on exclusive mode for ONE node
e. Create the new disk or Use existing disk
f. Restore the voting disk to newly created node
g. Stop the CRS which is started on exclusive mode
h. Start the CRS on both node
i. Start the database and make sure all the instance are up
j. Verify the cluster service and make sure all good!

26. What is OCR and what are the information it stored?

OCR is a file that manages the cluster and Real Application Clusters (RAC) database configuration information.

The OCR serves as the central repository for Cluster Ready Services (CRS); storing the metadata, configuration, and state information for all cluster resources defined in Clusterware. The OCR always retains the latest three backup copies of the OCR, which are four hours old, one day old, and one week old.

OCR has the following information:

- Node membership information, including which nodes are part of the cluster
- Software current active version
- Location of the voting disk
- Server pools

Location of OCR can be found as below:

ocrcheck or cat /etc/oracle/ocr.loc

27. Which commands can be used to check OCR health and backups?

ocrcheck
ocrconfig –showbackup auto -> To see the backups
ocrconfig –showbackup manual -> To see the manual backups
ocrconfig –restore <file_name> -> to restore OCR from the backup

28. What is OLR and use of it?

Oracle Local Registry contains information specific to the node in the cluster. This is the first file that is accessed to startup the cluster when OCR stored in the ASM.

OCR should be accessible to find out the resources which need to be started on a node. If OCR is on ASM, it can't be read until ASM (which itself is a resource for the node and this information is stored in OCR) is up.

To resolve this problem, information about the resources which need to be started on a node is stored in an operating system file which is called Oracle Local Registry or OLR. Since OLR is an operating system file, it can be accessed by various processes on the node for read/write irrespective of the status of the Clusterware (up/down). Hence, when a node joins the cluster, OLR on that node is read, various resources, including ASM are started on the node. Once ASM is up, OCR is accessible and is used to manage all the Clusterware resources. If OLR is missing or corrupted, Clusterware can't be started on that node!

Location of OLR can be found in: /etc/oracle/olr.loc or ocrcheck –local
Default location of OLR: $GRID_HOME/cdata/*olr
To see the backups of OLR: ocrcheck –local -showbackup
To restore OLR: ocrconfig -local -restore file_name
To check integrity of OLR: cluvfy comp olr

29. Explain the process Oracle RAC database startup?

1. **Operating System Initialization**

   When a system starts, it reads /etc/init.d file and triggers OHASD (Oracle High Availability Service Daemon)

2. **OHASD Initialization**

   OHASD is responsible to initiate the Oracle clusterware.

   OHASD reads the OLR stored locally and use the necessary information to start the clusterware.

3. **GPNPD and CSSD Initialization**

   OHASD starts the GPnP daemon and CSS (Cluster Synchronization Service) daemon and then CSS daemon access GPnP profile and find the below information.

   ASM Diskgroup Discovery String

   ASM SPFILE location (Diskgroup name)

   ASM Diskgroup containing the Voting Files

4. **Voting disk and CSSD**

   CSSD will locate the voting files stored in ASM disk and completes the initialization and decides to start/join the cluster.

5. **ASM Instance initialization**

   OHASD starts an ASM instance, by locating the SPFILE usually stored in the diskgroup The ASM instance mounts its Diskgroups, providing access to the Clusterware's Oracle Cluster Registry (OCR).

6. **CRSD Initialization**

   OHASD starts the Cluster Ready Services Daemon (CRSD), which accesses the OCR in an ASM Diskgroup. With CRSD running, Clusterware completes its initialization and brings up other essential services.

   In short, below are the 5 files accessed to start the Oracle RAC.

FILE 1 : OLR (ORACLE LOCAL REGISTRY) ——————————-> OHASD Process
FILE 2 : GPNP PROFILE (GRID PLUG AND PLAY) ——————————> GPNPD process
FILE 3 : VOTING DISK ——————————————————————-> CSSD Process
FILE 4 : ASM SPFILE ——————————————————————> OHASD Process
FILE 5 : OCR (ORACLE CLUSTER REGISTRY) ——————————> CRSD Process

30. What are the RAC background processes?

a. **ACMS - Atomic Control file to Memory Service (ACMS)**

   Ensuring a distributed SGA memory update is either globally committed on success or globally aborted if a failure occurs.

b. **GTX0 - Global Transaction Process**

   The process provides transparent support for XA global transactions in a RAC environment. The database auto tunes the number of these processes based on the workload of XA global transactions.

c. **LMON - Global Enqueue Service Monitor**

   This process monitors global Enqueue and resources across the cluster and performs global Enqueue recovery operations. This is called as Global Enqueue Service Monitor.

d. **LMS - Global Cache Service Process**

   This process is called as Global Cache service process. This process maintains statuses of data files and each cached block by recording information in a Global Resource Directory (GRD).This process also controls the flow of messages to remote instances and manages global data block access and transmits block images between the buffer caches of different instances. This processing is a part of cache fusion feature.

e. **LMD - Global Enqueue Service Daemon**

   This process is called as global Enqueue service daemon. This process manages incoming remote resource requests within each instance.

f. LCK0 - Instance Enqueue Process

This process is called as Instance Enqueue process. This process manages non-cache fusion resource requests such as library and row cache requests.

g. RMSn - Oracle RAC Management Processes (RMSn)

This process is called as Oracle RAC management process. These processes perform manageability tasks for Oracle RAC. Tasks include creation of resources related Oracle RAC when new instances are added to the cluster.

h. RSMN - Remote Slave Monitor

This process is called as Remote Slave Monitor. This process manages background slave process creation and communication on remote instances. This is a background slave process. This process performs tasks on behalf of coordinating process running in another instance.

31. What are the role of CSSD, CRSD, CTSSD, EVMD, and GPNPD?

CSSD (Cluster Synchronization Service Daemon)

It manages the cluster configuration like, which nodes are part of cluster etc. . . . When a node is added or deleted, it inform the same about this other nodes. It is also responsible for node eviction if situation occurs. It has 3 processes

1. CSS daemon (ocssd),
2. CSS Agent (cssdagent), the cssdagent process monitors the cluster and provides input/output fencing.
3. CSS Monitor (cssdmonitor) – Monitors internode cluster health

CRSD (Cluster ready service daemon)

It manages the cluster resources, based on OCR information. It includes start, stop and failover or resource. It monitors database instance, asm instance, listeners, services and etc on and automatically restarts them when failure occurs.

CTSSD (Cluster Time Synchronization Service Daemon)

Provides time management for cluster service. If NTP is running on server, then CTSS runs in observer mode.

EVMD (Event Monitor)

Event Manger, Is a background process that publishes Oracle Clusterware events and manages message flow between the nodes and logs relevant information to log file.

GPnPD (Grid plug and play service daemon)

GPNPD provides access to the Grid Plug and Play profile, and coordinates updates to the profile among the nodes of the cluster to ensure that all of the nodes node have the most recent profile.

32. How oracle RAC manages the instance recovery in case of crash?

1. Normal RAC operation, all nodes are available.
2. One or more RAC instances fail.
3. Node failure is detected.
4. Global Cache Service (GCS) reconfigures to distribute resource management to the surviving instances.
5. The SMON process in the instance that first discovers the failed instance(s) reads the failed instance(s) redo logs to determine which blocks have to be recovered.
6. SMON issues requests for all of the blocks it needs to recover. Once all blocks are made available to the SMON process doing the recovery, all other database blocks are available for normal processing.
7. Oracle performs roll forward recovery against the blocks, applying all redo log recorded transactions
8. Once redo transactions are applied, all undo records are applied, which eliminates non-committed transactions.
9. Database is now fully available to surviving nodes.

33. What is cache fusion?

Cache Fusion, nothing but the transfer of blocks between two instances in Oracle Real Application Clusters (RAC), is the main and most important feature of RAC.

According to Oracle in Oracle RAC Cache Fusion, each "instance [in a] RAC cluster has its local buffer cache where it does cache functionality. But when multiple users are connected to different nodes, users often need to access or lock a data block owned by another instance.

In such cases, the requesting instance requests a holding instance for that data block and accesses it through an interconnected mechanism. This concept is known as *Cache Fusion*.

Three modes of locks.

1. Null (N) mode - Null mode is usually held as a placeholder.

2. Shared (S) mode - In this mode, a data block is not modified by another session but will allow concurrent shared access.

3. Exclusive (X) mode – This mode allows the holder to have exclusive access, other processes cannot write to the block.

Cache fusion has 3 different scenarios.

1. Read – read scenario

   Currently a block is owned by instance A for reading. Instance B requests a read to the same block to GCS (Global Cache Service – who is responsible to transfer blocks between the instances). Now a shared lock is acquired and GCS requests instance A to ship the requested block to instance B.

2. Read – write scenario

   Instance A is updating a block hence an exclusive lock is acquired on the block. Now instance B sends a read request to GCS. Now GCS checked and found that an exclusive lock is acquired on the block by instance A and inform instance A to release the block for read. Instance A will create a CR (Consistent Read – CR image is created using UNDO data, i.e the value before the block is modified) image and inform the GCS to ship the block to instance B.

3. Write – write scenario

   Instance A has acquired an exclusive lock in a block and instance B requests the same block to GCS. When GCS notifies instance A, it will create a PI (Past Image of the modified block created using redo entries) and notifies GCP to send the block to instance B.

34. What is Consistent Read Image and Past Image?

Consistent Read

A consistent read is needed when a particular block is being accessed/modified by transaction [A1], and at the same time another transaction [A2] tries to access/read the block. If [A1] has not been committed, [A2] needs a consistent read [(non-modified block)] copy of the block to move ahead.

A CR copy is created using the UNDO data for that block."

Past Image

Consider an instance holding an exclusive lock on a data block for updates. If some other instance in the RAC needs the block, the holding instance can send the block to the requesting instance (instead of writing it to disk) by keeping a PI (Past Image) of the block in its buffer cache. PI is the copy of the data block before the block is written to the disk.

35. What is GCS (Global Cache Service) and GES (Global Enqueue Service)?

Global Cache Service (LMSn)

Service responsible to ship the blocks between instances by retaining block copies in memory.

Each such copy is called a past image (PI). It is also possible to have more than one PI of the data block depending on how many times the block was requested in [the] dirty stage."

Note: If you want to read a data block, it must be read in [a] consistent state. You are not allowed to read the changes made by others.

Global Enqueue Service (LMD)

The global Enqueue service (GES) tracks the status of all Oracle enqueuing mechanisms.

GES controls concurrency control on below

1. Transaction Locks
2. Library Cache Locks
3. Dictionary Cache Locks
4. Table Locks

36. What is GRD (Global Resource Directory)?

GRD stands for Global Resource Directory. The GES and GCS maintains records of the statuses of each datafiles and each cached block using global resource directory. This process is referred to as cache fusion and helps in data integrity.

37. **What is gc block lost?**

It indicates issue with interconnect. If a requested block is not received by the instance in 0.5 seconds, the block is considered to be lost.

38. What is gc buffer busy wait?

A session is trying to access a buffer in buffer cache, but that particular buffer is currently busy with global cache operation.

So during that time gc buffer busy wait will happen.

Example –

Let's say session A want to access block id 100, but currently that block is in buffer cache of session B. So session a requested session B LMS process to transfer the block.
While transfer is going on, session B also tried to access that block. But as already that block/buffer is already busy in global cache operation. Session B has to wait with wait event, gc buffer busy wait.
Reasons – Concurrency related, Right hand index growth.

Other reason might be lack of CPU, slow interconnect.

39. What is split brain syndrome?

Split brain syndrome occurs when the Oracle RAC nodes are unable to communicate with each other via private interconnect, but the communication between client and RAC node is maintained. This can cause data Integrity issues when the same block is read or updated by two nodes and changes done from one node are overwritten by the other node because the block being changed is not locked.

40. What is node eviction and when the node eviction happens?

Node eviction can be defined as the situation when a network heartbeat is failed between the nodes failed or disk heartbeat to vote disk is failed, cssd process removes the server from the cluster and rebooted automatically and try to join the cluster.

ocssd.bin is responsible to ensure the disk heartbeat as well as the network heartbeat.

There's a maximum delay in both heartbeats, the delay of network heartbeat is called MC (Misscount), and the disk heartbeat delay is called IOT (I/O Timeout). These 2 parameters are in seconds, by default Misscount < Disktimeout.

[grid@Linux-01 ~]$ crsctl get css misscount
CRS-4678: Successful get misscount 30 for Cluster Synchronization Services.

[grid@Linux-01 ~]$ crsctl get css disktimeout
CRS-4678: Successful get disktimeout 200 for Cluster Synchronization Services.

- Network disruption, latency, or missing network heartbeats.
- Delayed or missing disk heartbeats
- Slow interconnect or failures

- Unable to read/write or access the majority of the voting disks (files)
- Known Oracle Clusterware bugs
- Lack of sufficient resource (CPU/memory starvation) availability on the node for OS scheduling by key CRS daemon processes
- Manual termination of the critical cluster stack daemon background processes (css, cssdagent, cssdmonitor)
- No space left on the device for the GI or /var file system
- Sudden death or hang of CSSD process
- ORAAGENT/ORAROOTAGENT excessive resource (CPU, MEMORY, SWAP) consumption resulting in node eviction on specific OS platforms.

41. How Oracle decides which node to be evicted if private interconnect between nodes fails?

In 3 node RAC

Let's take that, there are nodes A, B, C are in the cluster. If network heartbeat of node A failed. Node B and Node C won't be able to ping to node A, But B and C can communicate between each other. So B and C will have 2 votes (one more self ping and other for ping to other node).

But A will have only one vote (i.e for the self ping). So A has less vote, oracle decides that A needs to evict.

In 2 node RAC

Let's say the nodes are A, B. If network heartbeat fails, then A and B won't be able to ping each other. So both A, B will have one vote each. So which node to be evicted?? Here quorum disk comes into play. This **quorum disk (voting disk)** also represents one vote. So both A and B will try to acquire that vote. Whoever acquire that quorum, gets 2 votes and stay in the cluster and other one gets evicted.

42. What happens when both instances work independently?

Same data block might get read/write by both the instances server processes. In both scenarios data integrity issue might occur as block changes by any instances will not get locked and recorded by other instances.

43. What is the solution to split-brain syndrome or who will decide which node to survive in the cluster?

Voting disk is the solution to split-brain syndrome. Voting disk will decide which node to survive and which one to be evicted.

44. What is I/O fencing in Oracle RAC?

When split-brain syndrome occurs, the failed node is prevented from accessing all the shared disk. This method is known as I/O fencing or Disk fencing or Failure fencing.

45. Can we install 12c database on 11g clusterware?

NO, clusterware version can be same or higher than the database version.

46. What are the software stacks in oracle clusterware?

From 11g onwards, lower stack is OHAS (Oracle High Availability Service) managed by ohasd and upper stack is CRS (Cluster Ready Service) managed by crsd.

47. Which are the commands to check cluster status in Oracle RAC?

crsctl stat res -t

crsctl check crs

crsctl stat res -t –init

crsctl check cluster –all

48. What checks should be done if CSSD is not coming up?

1. Check if voting disk is accessible.
2. If there's any issues with the private interconnect
3. If oracssd resource auto_start parameter is set to NEVER

49. What is Nodeapps in oracle RAC?

Nodeapps are standard set of oracle application services which are started automatically for RAC.

Node apps Include: VIP, network, adminhelper, ONS

50. What are the ways to identify master node in the Oracle RAC?

1. From occsd and crsd log files.
2. Query the view v$ges_resource
3. The node that store the OCR backup is the master node. (ocrconfig –showbackup)

51. Which files of Oracle RAC should reside in the shared storage?

All data files, control files, SPFIles, redo log files must reside on cluster-aware shred storage.

52. Which are the shared storage solutions that supports Oracle RAC?

ASM (automatic storage management), raw disk devices, network file system (NFS), OCFS2 and OCFS (Oracle Cluster Filesystems).

53. What is rebootless node fencing?

Rebootless node fencing in Oracle Real Application Clusters (RAC) is a feature that attempts to gracefully stop the Grid Infrastructure (GI) on an evicted node instead of rebooting it.

Below are the sequence of steps:

- First Clusterware finds which node to be evicted
- Then I/O generating processes will be killed on the problematic node.
- Clusterware resources will be stopped on the problematic node
- OHASD process would be running, will try continuously to start CRS, till issue is resolved.

54. What is recovery buddy feature in Oracle 19c?

Usually when instance is crashed in RAC, then one node is elected among the surviving nodes, for doing the recovery. And that elected node will read the redo logs of the crashed instance and do the recovery.

However in 19c, one instance will recovery buddy of another instance. Like.

Instance A is recovery buddy of instance B.

Instance B is recovery buddy of instance C.

Instance C is recovery buddy of instance A.

And this buddy instance will the track the block/redo changes of the mapped instance and keep them in its sga (in hash table).

So recovery buddy features helps in reducing the recovery time (as it eliminates the elect and redo read phase).

55. How can we improve global cache performance?

We can increase the number of LMS processes, by increasing gc_server_process.

We can set "_high_priority_processes"="LMS*|LGWR*"

56. What are the log file locations?

CRS: $ORACLE_BASE/diag/crs/<hostname>/crs/trace/alert.log

ASM: $ORACLE_BASE/diag/+asm/<ASM_SID>/trace/alert_<ASM_SID>.log

RDBMS-Instances: $ORACLE_BASE/rdbms/<DB_NAME>/<DB_SID>/trace/alert_<DB_SID>.log

PDBs: There is not specific log for PDBS. Please see the alert.log of the corresponding RDBMS instances

57. What is GIMR (Grid Infrastructure Management Repository)?

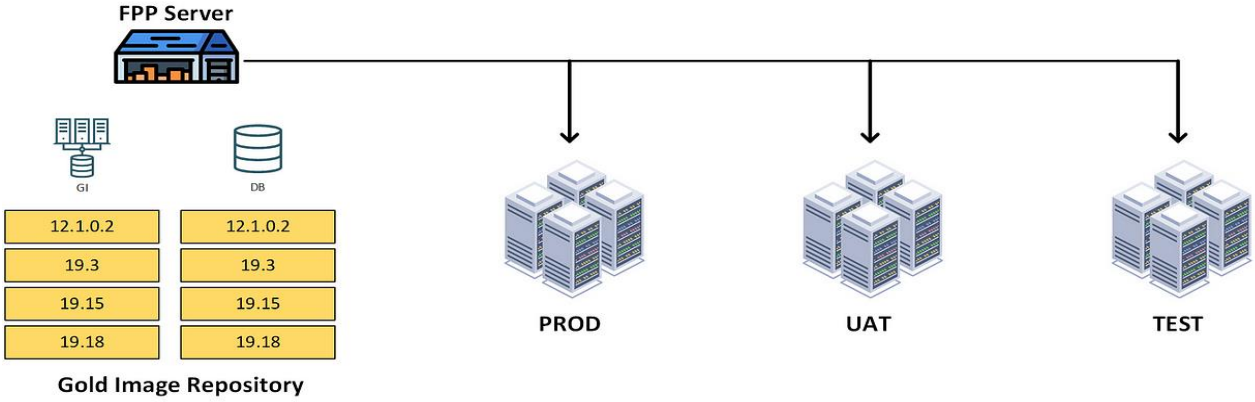
The Grid Infrastructure Management Repository (GIMR), or the Management Database (MGMTDB) is a multitenant database with a pluggable database (PDB) for the GIMR of each cluster. The GIMR stores the following information about the cluster:

- Real time performance data that Cluster Health Monitor collects
- Fault, diagnosis, and metric data that Cluster Health Advisor collects
- Cluster-wide events about all resources that Oracle Clusterware collects
- Workload performance and CPU architecture data that Quality of Service Management (QoS) collects
- Metadata required for Oracle Fleet Patching and Provisioning

58. What is Oracle fleet patching and provisioning?

Oracle Fleet patching and provisioning is a solution to the pain of upgrading and patching Maximum Availability databases.

Imagine, if we have 10 application and each application requires TEST, DEV, UAT, and PROD. Hence patching 40 databases is a tedious task for a DBA. Oracle Fleet patching and provisioning offers a solution to this problem by executing a single command.



Oracle FPP provides a tool capable of patching and upgrading Oracle Grid Infrastructure and Oracle Database using gold image. Gold image is a compressed copy of a software-only Oracle Database home or Grid Infrastructure home.

Oracle FPP server acts like a warehouse of various version of gold image that can be used to patching and upgrading.

59. What is a rolling upgrade?

A patch is considered a rolling if it is can be applied to the cluster binaries without having to shutting down the database in a RAC environment. All nodes in the cluster are patched in a rolling manner, one by one, with only the node which is being patched unavailable while all other instance open.

60. What are the high level steps to install and configure Oracle RAC in Linux?

1. Make operating systems ready (Node 1 and Node 2)

   Create users and groups, host file entries, and create directories for installation, IP assignments, DNS configuration, NTP configuration, shared disks, install cvuqdisk package, copy the grid software and unzip)
2. Configure Oracle ASM
   - Format the disks (fdisk /dev/sdb)
   - Configure oracleasm utility (/usr/sbin/oracleasm configure –i)
   - Initialize the asmlib with oracleasm init command (/usr/sbin/oracleasm init)
   - Create disk
     - oracleasm createdisk ASM_DISK1 /dev/sdb1
     - oracleasm scandisks

- oracleasm listdisks
- Scan the disks in second node
  - /usr/sbin/oracleasm configure –i
  - oracleasm scandisks
  - oracleasm listdisks

3. Pre-check for RAC Setup

   Execute cluvfy statement:

   ./runcluvfy.sh stage -pre crsinst -n dbwr1, dbwr2 -verbose

4. Install and Configure Oracle Grid Infrastructure for a Cluster

   ./gridSetup.sh

5. Oracle 19c Database Installation

   ./runInstaller – install software only

   Create disk group for database creation (use asmca)

   Create database using dbca

6. Post-Check For RAC Setup

   srvctl config database –d db_name

61. What are the high level steps to convert a non-RAC database to RAC database?

Two phase approach can be followed to convert a database from non-RAC to RAC.

1. Duplicate single instance non-ASM database to ASM using RMAN

2. Manually convert single instance database to RAC

Prerequisites

- Install and configure Clusterware
- Install and configure Automatic Storage Management (ASM)
- Install Oracle Database Software
- Standalone database and RAC must be on same Database version

1. Estimate used space for Non-RAC database

2. Create password file and init.ora file for RAC environment

   Initfile contents (basic)

```
db_name ="ORADB"
instance_name =ORADB
control_files =‘+DATA’, ‘+FLASH’
db_create_file_dest='+DATA'
db_recovery_file_dest='+FLASH'
db_create_online_dest_1=’+DATA’
db_create_online_dest_2=’+FLASH’
log_archive_dest_1='LOCATION=+FLASH’
```

3. Configure Oracle Listener and tnsnames.ora file

4. Duplicate the database to RAC server using RMAN

62. What are the high-level steps to add a node in Oracle RAC environment?

1. Add host entries, create users and groups and install ASM libraries, configure ASM, configure password less SSH, NTP as in the existing nodes.

2. Verify cluster utility.

./cluvfy comp peer -n racpb3 -refnode racpb1 -r 11gr2

./cluvfy stage -pre nodeadd -n racpb3 -fixup -verbose > /home/oracle/cluvfy_pre_nodeadd.txt

3. Add the node from Node 1.

./addnode.sh -silent "CLUSTER_NEW_NODES= {racpb3}"
"CLUSTER_NEW_VIRTUAL_HOSTNAMES= {racvr3}"

63. What are the high-level steps to apply patches in rolling manner in Oracle RAC?

1. Perform the patching pre-checks

   srvctl config database –verbose

   crsctl query crs softwareversion –all

   opatch version

   opatch lsinventory -detail -oh $GRID_HOME

   opatch prereq CheckConflictAgainstOHWithDetail -phBaseDir /u01/30135696/29401763

2. On Node 1: apply the patch using opatchauto command.

   opatchauto apply /u01/30135696

   opatchauto will bring down services on the patching node and bring up the services after patching

3. On Node 2: apply the patch using opatchauto command.

   opatchauto apply /u01/30135696

4. Execute datapatch and run utlrp if any objects are invalid.

   nohup ./datapatch -verbose &

5. Check the patching status using the following queries.

   SELECT description, ACTION, STATUS, TARGET_BUILD_TIMESTAMP from dba_registry_sqlpatch;

   crsctl query crs releasepatch