Top 51 RAC Interview Questions and Answer

Make sure to refresh your knowledge by reviewing this list of RAC Interview Questions.

1)What is cache fusion?

In a RAC environment, it is the combining of data blocks, which are shipped across the interconnect from remote database caches (SGA) to the local node, in order to fulfill the requirements for a transaction (DML, Query of Data Dictionary).

Cache fusion is nothing but a mapping of remote memory of oracle buffers, which is shared between the caches participating nodes in the cluster. It is very easy to gain the block image from the instance that contain the block in its SGA instead of reading from the disk, this happens when the block of data is read from data file by an instance in the cluster and when another instance require the same block.

2)What is split brain?

When database nodes in a cluster are unable to communicate with each other, they may continue to process and modify the data blocks independently. If the

same block is modified by more than one instance, synchronization/locking of the data blocks does not take place and blocks may be overwritten by others in the cluster. This state is called split brain.

3)What is the difference between Crash recovery and Instance recovery?

When an instance crashes in a single node database on startup a crash recovery takes place. In a RAC enviornment the same recovery for an instance is performed by the surviving nodes called Instance recovery.

4)What is the interconnect used for?

It is a private network which is used to ship data blocks from one instance to another for cache fusion. The physical data blocks as well as data dictionary blocks are shared across this interconnect.

5)How do you determine what protocol is being used for Interconnect traffic?

One of the ways is to look at the database alert log for the time period when the database was started up.

6)What methods are available to keep the time synchronized on all nodes in the cluster?

Either the Network Time Protocol(NTP) can be configured or in 11gr2, Cluster Time Synchronization Service (CTSS) can be used.

7)What files components in RAC must reside on shared storage?

Spfiles, ControlFiles, Datafiles and Redolog files should be created on shared storage.

8)Where does the Clusterware write when there is a network or Storage missed heartbeat?

The network ping failure is written in $CRS_HOME/log

9)How do you find out what OCR backups are available?

The ocrconfig -showbackup can be run to find out the automatic and manually run backups.

10)If your OCR is corrupted what options do have to resolve this?

You can use either the logical or the physical OCR backup copy to restore the Repository.

11)How do you find out what object has its blocks being shipped across the instance the most?

You can use the dba_hist_seg_stats.

12)What is a VIP in RAC use for?

The VIP is an alternate Virtual IP address assigned to each node in a cluster. During a node failure the VIP of the failed node moves to the surviving node and relays to the application that the node has gone down. Without VIP, the application will wait for TCP timeout and then find out that the session is no longer live due to the failure.

13)How do we know which database instances are part of a RAC cluster?

You can query the V$ACTIVE_INSTANCES view to determine the member instances of the RAC cluster.

14)What is OCLUMON used for in a cluster environment?

The Cluster Health Monitor (CHM) stores operating system metrics in the CHM repository for all nodes in a RAC cluster. It stores information on CPU, memory, process, network and other OS data, This information can later be retrieved and used to troubleshoot and identify any cluster related issues. It is a default component of the 11gr2 grid install. The data is stored in the master repository and replicated to a standby repository on a different node.

15) What would be the possible performance impact in a cluster if a less powerful node (e.g. slower CPU's) is added to the cluster?

All processing will show down to the CPU speed of the slowest server.

16) What is the purpose of OLR?

Oracle Local repository contains information that allows the cluster processes to be started up with the OCR being in the ASM storage ssytem. Since the ASM file system is unavailable until the Grid processes are started up a local copy of the contents of the OCR is required which is stored in the OLR.

17) What is the default memory allocation for ASM?

In 10g the default SGA size is 1G in 11g it is set to 256M and in 12c ASM it is set back to 1G.

18) How do you backup ASM Metadata?

You can use md_backup to restore the ASM diskgroup configuration in-case of ASM diskgroup storage loss.

19) What files can be stored in the ASM diskgroup?
In 11g the following files can be stored in ASM diskgroups.

- Datafiles
- Redo logfiles
- Spfiles

   In 12c the files below can also new be stored in the ASM Diskgroup

- Password file

20) What it the ASM POWER_LIMIT?

This is the parameter which controls the number of Allocation units the ASM instance will try to rebalance at any given time. In ASM versions less than 11.2.0.3 the default value is 11 however it has been changed to unlimited in later versions.

21) 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.

22)What are some of the RAC specific parameters?

Some of the RAC parameters are:

* CLUSTER_DATABASE
* CLUSTER_DATABASE_INSTANCE
* INSTANCE_TYPE (RDBMS or ASM)
* ACTIVE_INSTANCE_COUNT
* UNDO_MANAGEMENT

23)What is the future of the Oracle Grid?

The Grid software is becoming more and more capable of not just supporting HA for Oracle Databases but also other applications including Oracle's applications. With 12c there are more features and functionality built-in and it is easier to deploy these pre-built solutions, available for common Oracle applications.

24)What components of the Grid should I back up?

The backups should include OLR, OCR and ASM Metadata.

25)Is there an easy way to verify the inventory for all remote nodes?

You can run the opatch lsinventory -all_nodes command from a single node to look at the inventory details for all nodes in the cluster.

26)What are the components of clusterware?

Oracle cluster registry (OCR): It contains all information about instances, services, state information, cluster configuration, nodes and ASM storage if needed. The OCR should occupy on a shared disk, which is accessible by all the nodes in your cluster. In OCR, the daemon OCSSd is used to manage the configuration and in the registry, it maintains the changes to the cluster.
Voting Disk: It helps to verify, if a node has failed, which means it got separated from the majority, then it is rebooted forcibly and after rebooting, it is added again to the surviving nodes of cluster. The Oracle RAC uses it to maintain the membership of cluster.

27)What is FAN?

FAN stands for fast application notification, which is connected to the events containing services, nodes and instances. In order to describe the other processes about the service level information and configuration which contains the changes of the service status like UP or DOWN events, Oracle RAC 12c uses this notification mechanism. Using FAN events, the application gives response and can take immediate actions.

28)What is SCAN?

SCAN stands for Single Client Access Name is a feature of new Oracle RAC 11g release 2 which provides one name for clients to access an Oracle Database cluster. The benefit to the SCAN user is that, there is no need to change if you remove or add nodes in the cluster.

29)What is hangcheck timer?

The hangcheck timer is used to check the health of the system regularly. The node is restarted automatically, when the system stops or hangs.

There are two key parameters. They are:

Hangcheck margin- this shows that how much delay can be permitted before the reset of the RAC node is done by the hangcheck timer.
Hangcheck Tick: It is defined as the time period between system health checks. 60 seconds is the default time, but Oracle recommends it to be 30 seconds.

30)What is GRD?

GRD is the Global Resource Directory. The GRD is used by the GES and GCS to maintain the records of each cached block and each datafile. This process is known as cache fusion and can be used in data integrity.

31)What is OCR file?

It is a RAC configuration information repository, which maintains the information about the instance node mapping and cluster node. It also maintains information about the profiles of oracle Clusterware resource for customed applications. It manages the configuration information related to any cluster database in the cluster. It is necessary for the OCR to reside on a shared disk, which is accessible by all of the cluster nodes. The command daemon OCSSD maintains the configuration information in OCR and manages the changes to cluster within the registry.

32)What is a raw partition?

It is the part of the physical disk, which is accessed on the lowest level. When an addition partition is created, raw partition is created and without any formatting, a logical partitions are assigned to it. It is called cooked partition, once the formatting is completed.

33)What is the use of VIP?

When the node fails, then the VIP address of that node fails over to the other node on which it cannot accept Oracle connections but not TCP connections.

34)What is load balancing Advisory?

To check the workload across resources in the balancing of application, the load balancing advisory is provided.

35)What options are required to resolve OCR, if it is corrupted?

The backup copy of either physical or logical OCR copy is used to restore the repository.

36) What is Voting file/disk and how many files should be there?
Voting Disk File is a file on the shared cluster system or a shared raw device file. Oracle Clusterware uses the voting disk to determine which instances are members of a cluster. Voting disk is akin to the quorum disk, which helps to avoid the split-brain syndrome. Oracle RAC uses the voting disk to manage cluster membership by way of a health check and arbitrates cluster ownership among the instances in case of network failures. The voting disk must reside on shared disk.

37) How is possible to install a RAC if we don't have a CFS?

This is possible by using a raw device.

38) What is a raw device?

A raw device is a disk drive that does not yet have a file system set up. Raw devices are used for Real Application Clusters since they enable the sharing of disks.

39) When to use CFS(Cluster File System) over raw?

A CFS offers:
– Simpler management
– Use of Oracle Managed Files with RAC
– Single Oracle Software installation
– Autoextend enabled on Oracle datafiles
– Uniform accessibility to archive logs in case of physical node failure
– With Oracle_Home on CFS, when you apply Oracle patches CFS guarantees that the updated Oracle_Home is visible to all nodes in the cluster.
Note: This option is very dependent on the availability of a CFS on your platform.

40) When to use raw over CFS?

– Always when CFS is not available or not supported by Oracle.
– The performance is very, very important: Raw devices offer best performance without any intermediate layer between Oracle and the disk.
Note: Autoextend fails on raw devices if the space is exhausted. However the space could be added online if needed.

41) Why we need to have configured SSH or RSH on the RAC nodes?

SSH (Secure Shell,10g+) or RSH (Remote Shell, 9i+) allows "oracle" UNIX account connecting to another RAC node and copy/ run commands as the local "oracle" UNIX account.

42) Is the SSH, RSH needed for normal RAC operations?

No. SSH or RSH are needed only for RAC, patch set installation and clustered database creation.

43) Does Real Application Clusters support heterogeneous platforms?

The Real Application Clusters do not support heterogeneous platforms in the same cluster.

44) What is the Cluster Verification Utiltiy (cluvfy)?

The Cluster Verification Utility (CVU) is a validation tool that you can use to check all the important components that need to be verified at different stages of deployment in a RAC environment.

45) What versions of the database can I use the cluster verification utility (cluvfy) with?

The cluster verification utility is release with Oracle Database 10g Release 2 but can also be used with Oracle Database 10g Release 1.

46) How to take backup and recover of OCR file?
Backup :
#ocrconfig -manualbackup
#ocrconfig -export file_name.dmp
#ocrdump -backupfile my_file
$cp -p -R /u01/app/crs/cdata /u02/crs_backup/ocrbackup/RAC1

Recover:
#ocrconfig -restore backup_file.ocr
#ocrconfig -import file_name.dmp

47) What is local OCR?

/etc/oracle/local.ocr
/var/opt/oracle/local.ocr

48) How do you backup voting disk ? and how do I identify the voting disk location?

Backup voting disk command:

#dd if=voting_disk_name of=backup_file_name

Identify the voting disk location:

#crsctl query css votedisk

49) How do I identify the OCR file location?

check /var/opt/oracle/ocr.loc or /etc/ocr.loc (depends upon platform)
or
#ocrcheck

50) What is the functionality and features provided by Oracle Clusterware?

- Oracle Clusterware is a portable solution that comes up with the Oracle database to be used by the application that is being integrated and designed.

- It provides a complete solution of all the clustering problems and supports all the applications that use the functionality of Oracle database.

- It helps in implementation of Oracle RAC and support the clustered database as well that allow more and more applications to run.

- It manages all the Oracle RAC database integrations on the server and system with their statuses and provides an automatic recovery in case of any failure occurs.

- It provides high availability APIs that can be used with the applications and manage to provide a framework that can be integrated and used to develop more applications.

- It allows the user to use the registration services that allow starting, stopping and monitoring of the processes if any failure occurs.

51) Difference between crsctl and srvctl utility in RAC?

CRSCTL manages clusterware-related operations:
- Starting and stopping Oracle Clusterware
- Enabling and disabling Oracle Clusterware daemons
- Registering cluster resources
-

SRVCTL manages Oracle resource–related operations:
- Starting and stopping database instances and services
- Also from 11gR2 manages the cluster resources like network, vip, disks etc

CRSCTL and SRVCTL:

crsctl command is used to manage the elements of the **clusterware** (crs,cssd, OCR,voting disk etc.) while **srvctl** is used to manage the elements of the **cluster** (databases,instances,listeners, services etc) . For exemple, with **crsctl** you can tune the heartbeat of the cluster, while with **srvctl** you will set up the load balancing at the service level. Both command were introduced with Oracle 10g and have been improved since. There is sometimes some confusion among DBAs because both commands can be used to start the database, **crsctl** starting the whole clusterware + cluster, while**srvctl** is starting the other elements, such as database, listener, services, *but not the clusterware*. Somewhere in Oracle documentation 10g, it is written that Oracle corporation suggest you to use srvctl command to start the databases in the cluster.

What is srvctl?

It is the Server Control Utility, we use SRVCTL to start and stop the database and instances, manage configuration information, and to add, move or remove instances and services.

srvctl commands:

- To start a rac database
- To stop a rac database
- To check status and configurations
- To start and stop instances
- To start, stop and manage services

Start a rac database (order: nodeapps – asm – database)
srvctl start nodeapps -n *nodename*
srvctl start asm -n nodename
srvctl start database -d dbname
options are: srvctl start database -d dbname -o open | -o mount | -o nomount

Stop a rac database (order: database – asm – nodeapps)
srvctl stop database -d dbname -o immediate
options are: srvctl stop database -d dbname -o normal | -o transactional | -o immediate | -o abort
srvctl stop asm -n nodename
options are: srvctl stop asm -n nodename -o immediate
srvctl stop nodeapps -n nodename

To check status and configurations
Nodeapps:
srvctl status nodeapps -n nodename
srvctl config nodeapps -n nodename
ASM:
srvctl status asm -n nodename
srvctl config asm -n nodename
Database:
srvctl status database -d dbname
srvctl config database -d dbname (shows instances name, node and oracle home)
Instance:
srvctl status instance -d dbname -i instancename
Services:
srvctl status service -d dbname

To start and stop instances
srvctl start instance -d dbname -i instancename
srvctl stop instance -d dbname -i instancename

To start, stop and manage services
srvctl status service -d dbname
srvctl config service -d dbname
srvctl start service -d dbname -s servicename
srvctl stop service -d dbname -s servicename