Oracle RAC Interview Questions and Answers

Level 4 DBA - Senior Production Support and Troubleshooting Guide

Covers RAC architecture, Clusterware, OCR/voting disks, ASM, networking, services, cache fusion, GC waits, node eviction, patching, and real production scenarios.

These questions are suitable for a senior / Level 4 Oracle DBA, where the interviewer expects strong production troubleshooting, architecture knowledge, and real-time issue-handling experience.

1. Oracle RAC Architecture

1. What is Oracle RAC?

Oracle RAC, or Real Application Clusters, allows multiple Oracle instances running on different servers to access the same database.

```
Node 1 Instance  ---> Shared Database
Node 2 Instance  ---> Shared Database
Node 3 Instance  ---> Shared Database
```

RAC provides:

- High availability
- Scalability
- Load balancing
- Instance failover
- Service failover
- Rolling maintenance support

In RAC, each node has its own instance, memory, background processes, and redo thread, but all instances access the same database files on shared storage.

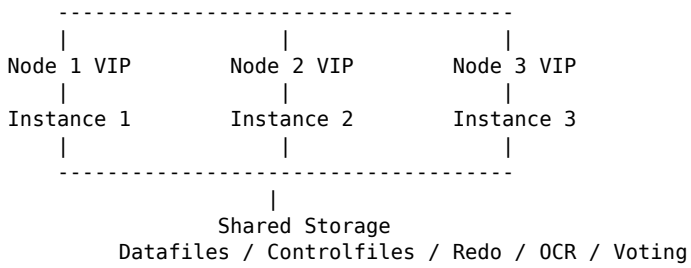
2. Explain Oracle RAC architecture.

Oracle RAC architecture contains:

Component	Purpose
Clusterware	Manages cluster nodes and resources
Voting Disk	Maintains node membership
OCR	Stores cluster configuration
ASM	Manages shared storage
Interconnect	Private network for cache fusion
Public Network	Client/application connectivity
SCAN	Single Client Access Name for connection load balancing
Local Listener	Listener running on each node
Remote Listener	Usually points to SCAN listeners
RAC Instances	Oracle instances on each node
Shared Database	Common datafiles, control files, redo logs

High-level diagram:

```
                    Application / Client
                             |
                             v
                     SCAN Name / SCAN VIPs
                             |
```



3. What is the difference between RAC database and single-instance database?

Single Instance	RAC
One instance accesses database	Multiple instances access same database
One server	Multiple servers
No cache fusion	Uses cache fusion
Instance failure causes DB outage unless failover configured	Surviving instances continue
Simpler architecture	More complex but highly available

4. What are the main background processes in RAC?

Important RAC-specific processes:

Process	Purpose
LMS	Global Cache Service process, transfers blocks
LMD	Global Enqueue Service daemon
LMON	Global Enqueue Service monitor
LCK	Lock process
DIAG	Diagnostic process
GCS	Global Cache Service
GES	Global Enqueue Service

Common Oracle processes like DBWn, LGWR, CKPT, SMON, and PMON also exist on each RAC instance.

2. Clusterware

5. What is Oracle Clusterware?

Oracle Clusterware is the software layer that manages the cluster.

It manages:

- Cluster membership
- Node monitoring
- Resource startup and shutdown
- VIPs
- SCAN listeners

- Database instances
- Services
- ASM
- Node eviction

Clusterware ensures that only healthy nodes remain part of the cluster.

6. What are the main Oracle Clusterware components?

Component	Purpose
OHASD	Starts Oracle High Availability services
CSSD	Manages cluster synchronization and voting disk access
CRSD	Manages cluster resources
EVMD	Event management daemon
CTSSD	Time synchronization
ASM	Shared storage management
GIPCD	Grid interprocess communication
GPNPD	Grid Plug and Play profile daemon
MDNSD	Multicast DNS daemon

7. What is OHASD?

OHASD means Oracle High Availability Services Daemon.

It is the first major Clusterware daemon started on a node. It starts lower-level cluster components such as CSSD, CRSD, EVMD, and ASM.

If OHASD is down, Clusterware resources cannot run properly on that node.

8. What is CSSD?

CSSD means Cluster Synchronization Services Daemon.

It manages cluster membership, node heartbeat, voting disk communication, and split-brain prevention.

If CSSD detects that a node is unhealthy or cannot communicate properly, node eviction may happen.

9. What is CRSD?

CRSD means Cluster Ready Services Daemon.

It manages high-level cluster resources such as database, instances, services, listeners, VIPs, SCAN VIPs, and ASM resources.

If CRSD is down, resources may not be managed correctly.

10. What is EVMD?

EVMD means Event Manager Daemon.

It publishes and handles cluster events such as node join, node leave, resource failure, VIP relocation, and service relocation.

3. OCR and Voting Disk

11. What is OCR?

OCR means Oracle Cluster Registry.

It stores cluster configuration information such as cluster nodes, database resources, ASM resources, services, VIPs, SCAN configuration, and resource profiles.

OCR is critical for Clusterware startup and management.

12. What is Voting Disk?

Voting Disk is used to determine cluster membership.

It helps decide which nodes are part of the cluster and prevents split-brain. Voting disk stores heartbeat information from cluster nodes.

13. Difference between OCR and Voting Disk?

OCR	Voting Disk
Stores cluster configuration	Stores node membership/heartbeat
Used by CRSD	Used by CSSD
Required for resource management	Required for cluster membership
Can be backed up automatically/manual	Usually placed in ASM diskgroup

14. How do you check OCR location?

```
ocrcheck
ocrconfig -showbackup
```

15. How do you check Voting Disk location?

```
crsctl query css votedisk
```

16. How do you backup OCR manually?

```
ocrconfig -manualbackup
ocrconfig -showbackup
```

17. What happens if OCR is corrupted?

Cluster resources may not start or may not be managed.

Recovery options:

- Stop Clusterware if required.
- Restore OCR from backup.
- Validate OCR.
- Restart Clusterware.

Command example:

```
ocrconfig -restore <backup_file>
```

18. What happens if Voting Disk is lost?

If the cluster loses access to enough voting disks, nodes may be evicted or the cluster may go down.

With normal redundancy, the cluster can tolerate some voting disk failures. The key point is that Clusterware must maintain voting disk quorum.

4. ASM in RAC

19. What is ASM?

ASM means Automatic Storage Management.

It manages Oracle database storage using disk groups. In RAC, ASM provides shared storage access for all nodes.

ASM stores datafiles, control files, redo logs, archive logs, OCR, voting disks, and SPFILEs.

20. What is ASM disk group redundancy?

Redundancy	Meaning
External	Storage provides redundancy
Normal	ASM two-way mirroring
High	ASM three-way mirroring

21. What is ASM rebalance?

ASM rebalance redistributes extents across disks after disk add, drop, or resize.

Check rebalance:

```
SELECT group_number, operation, state, power, actual, sofar, est_work, est_minutes
FROM v$asm_operation;
```

Increase rebalance power:

```
ALTER DISKGROUP DATA REBALANCE POWER 8;
```

22. ASM disk dropped accidentally. What will you do?

First check diskgroup status and redundancy.

```
SELECT name, state, type FROM v$asm_diskgroup;
SELECT name, path, mount_status, header_status, mode_status FROM v$asm_disk;
```

If the disk is temporarily unavailable, bring it back before rebalance completes.

If the disk is permanently failed:

- Confirm redundancy level.
- Replace failed disk.
- Add new disk to diskgroup.
- Monitor rebalance.
- Check alert logs.

Never force operations without understanding redundancy and data-loss risk.

5. RAC Networking

23. What are the networks used in RAC?

Network	Purpose
Public network	Client/application connectivity
Private network	Interconnect/cache fusion
VIP	Fast failover for client connections
SCAN network	Cluster-wide client access
ASM network	Optional dedicated ASM traffic in newer versions

24. What is private interconnect?

Private interconnect is the private network used by RAC nodes to communicate.

It is used for cache fusion, heartbeats, global cache messages, and instance coordination.

A slow or unstable interconnect can cause RAC performance issues, GC waits, node eviction, and split-brain risk.

25. What is SCAN?

SCAN means Single Client Access Name.

It provides a single name for clients to connect to RAC. SCAN resolves to multiple SCAN VIPs, usually three.

Benefits:

- Simplifies client connection string
- Enables connection load balancing
- Allows adding/removing nodes without changing client TNS

26. What is VIP in RAC?

VIP means Virtual IP.

Each node has a VIP. If a node fails, its VIP relocates to another node. This allows clients to receive fast connection failure instead of waiting for TCP timeout.

27. Difference between VIP and SCAN VIP?

VIP	SCAN VIP
Node-specific virtual IP	Cluster-wide virtual IP
One per node	Usually three per cluster
Used after client is redirected to node	Used as initial contact point
Relocates if node fails	Relocates among nodes if needed

28. What is local listener and remote listener?

Local listener runs on each node and listens for local instance connections.

Remote listener usually points to SCAN listeners.

```
LOCAL_LISTENER  = node VIP listener
REMOTE_LISTENER = SCAN name:port
```

Remote listener helps with service registration and load balancing.

6. Services in RAC

29. What are services in RAC?

Services are logical workload names used for application connections.

They allow workload separation, load balancing, failover, role-based services, application continuity, and instance affinity.

Example:

```
PAYROLL_SERVICE runs on instance 1 and 2
REPORT_SERVICE runs on instance 3
```

30. Why should applications connect using services instead of instance name?

Applications should use services because services support failover, load balancing, planned relocation, role-based configuration, and maintenance without changing connection strings.

Connecting directly to instance names is not recommended in RAC.

31. How do you create a service in RAC?

```
srvctl add service -db PRODDB -service APP_SVC -preferred PRODDB1,PRODDB2 -available PRODDB3
srvctl start service -db PRODDB -service APP_SVC
srvctl status service -db PRODDB
```

32. What is preferred and available instance?

Type	Meaning
Preferred Instance	Service normally runs here
Available Instance	Service can fail over here

33. What is service relocation?

Service relocation means moving a service from one instance to another.

```
srvctl relocate service -db PRODDB -service APP_SVC -oldinst PRODDB1 -newinst PRODDB2
```

It is used during maintenance or load balancing.

7. Cache Fusion and Global Cache

34. What is Cache Fusion?

Cache Fusion is the RAC mechanism that transfers data blocks between instances through the interconnect.

Instead of reading a block from disk, one instance can receive the block from another instance's buffer cache.

```
Instance 1 has block A in buffer cache.
Instance 2 needs block A.
Instance 1 sends block A to Instance 2 through interconnect.
```

This reduces disk I/O but depends heavily on a fast interconnect.

35. What is Global Cache Service?

Global Cache Service, or GCS, manages block transfers between RAC instances.

It tracks which instance has which block and in what mode. GCS is mainly handled by LMS processes.

36. What is Global Enqueue Service?

Global Enqueue Service, or GES, manages global locks and enqueues across RAC instances.

It coordinates access to resources such as library cache locks, DDL locks, transaction locks, and dictionary locks.

37. What are current block and CR block in RAC?

Block Type	Meaning
Current Block	Latest version of block, used for modification
CR Block	Consistent read version of block, used for queries

A SELECT may request a CR block. An UPDATE may request a current block.

8. RAC Wait Events

38. What are common RAC wait events?

Wait Event	Meaning
gc cr request	Waiting for consistent read block
gc current request	Waiting for current block
gc buffer busy acquire	Waiting for busy global cache block
gc buffer busy release	Waiting while releasing busy block
gc cr block busy	CR block busy
gc current block busy	Current block busy
gc cr block 2-way	CR block received from another instance
gc current block 2-way	Current block received from another instance
gc block lost	Block lost during interconnect transfer

39. What does high `gc cr request` mean?

It means sessions are waiting for consistent read blocks from another instance.

Possible causes:

- Heavy cross-instance reads
- Poor service placement
- Reporting queries reading blocks modified by another instance
- Hot objects accessed from multiple instances
- Slow interconnect

40. What does high `gc current request` mean?

It means sessions are waiting for current blocks from another instance for modification.

Possible causes:

- Same table/index blocks modified from multiple instances
- Hot block contention
- Sequence/index right-hand growth
- Poor workload partitioning
- Cross-instance DML

41. What causes high `gc buffer busy acquire`?

Common causes:

- Hot block contention
- Multiple instances modifying same block
- Right-hand index block contention

- Unpartitioned sequence-based inserts
- Poor application workload distribution
- Same rows updated by multiple instances
- Slow interconnect or slow LMS

42. How do you troubleshoot RAC GC waits?

Steps:

- Check AWR Global Report.
- Check AWR per-instance reports.
- Identify top GC wait event.
- Identify SQL causing GC waits.
- Identify hot segments.
- Check services placement.
- Check interconnect latency.
- Check LMS CPU usage.
- Check application workload distribution.

Useful query:

```
SELECT inst_id, event, total_waits, time_waited
FROM gv$system_event
WHERE event LIKE 'gc%'
ORDER BY time_waited DESC;
```

Hot segments:

```
SELECT owner, object_name, statistic_name, value
FROM gv$segment_statistics
WHERE statistic_name LIKE 'gc%'
ORDER BY value DESC;
```

9. Node Eviction

43. What is node eviction?

Node eviction means Clusterware removes a node from the cluster to protect database integrity.

It happens to prevent split-brain or because the node is unhealthy.

Common causes:

- Missed network heartbeat
- Missed disk heartbeat
- Private interconnect failure
- CPU starvation
- Memory starvation
- Storage freeze

- CSSD issue
- OS hang

44. What is split-brain?

Split-brain occurs when cluster nodes cannot communicate and each side thinks the other side is down.

This is dangerous because multiple nodes could try to modify shared database files independently.

Oracle uses voting disks and node eviction to prevent split-brain.

45. What are network heartbeat and disk heartbeat?

Heartbeat	Description
Network heartbeat	Communication between nodes over private interconnect
Disk heartbeat	Communication through voting disk

Both are used to determine cluster membership.

46. How do you troubleshoot node eviction?

Check:

- OS logs: /var/log/messages or syslog
- Cluster alert log
- CSSD logs
- CRS logs
- CHM/OSWatcher data
- Network logs
- Storage logs
- CPU/memory usage before eviction
- Interconnect packet loss
- Voting disk latency

Important locations:

```
$GRID_HOME/log/<hostname>/alert<hostname>.log
$GRID_HOME/log/<hostname>/cssd/
$GRID_HOME/log/<hostname>/crsd/
$GRID_HOME/log/<hostname>/ohasd/
```

Commands:

```
crsctl check cluster -all
crsctl stat res -t
olsnodes -n -s
```

47. What is CRS-1612 / CRS-1606?

These are cluster membership and heartbeat-related messages.

They usually indicate node communication issues, missed heartbeat, or eviction activity.

As a senior DBA, I would check CSSD logs, OSWatcher, network, storage latency, and CPU starvation during that timestamp.

10. Clusterware Commands

48. How do you check cluster status?

```
crsctl check cluster -all
```

49. How do you check cluster resources?

```
crsctl stat res -t
crsctl status resource -t
```

50. How do you check node list?

```
olsnodes -n -s
```

51. How do you stop/start Clusterware?

Stop on local node:

```
crsctl stop crs
```

Start on local node:

```
crsctl start crs
```

Stop cluster on all nodes:

```
crsctl stop cluster -all
```

Start cluster on all nodes:

```
crsctl start cluster -all
```

52. How do you check database status using SRVCTL?

```
srvctl status database -db PRODDB
srvctl start database -db PRODDB
srvctl stop database -db PRODDB
```

53. How do you check instance status?

```
srvctl status instance -db PRODDB -instance PRODDB1
srvctl start instance -db PRODDB -instance PRODDB1
srvctl stop instance -db PRODDB -instance PRODDB1
```

54. How do you check listener status?

```
srvctl status listener
srvctl status scan_listener
```

```
srvctl status scan
```

11. RAC Performance Tuning

55. How do you analyze RAC performance issue?

I first identify whether the issue is:

- SQL issue
- Instance-specific issue
- Cluster-wide issue
- Interconnect issue
- Storage issue
- Service placement issue
- Hot block issue

Then I check:

- AWR Global Report
- AWR per-instance report
- ASH by instance
- Top SQL
- GC wait events
- Interconnect statistics
- Segment statistics
- Services distribution

56. What is the difference between AWR Global Report and instance AWR report?

Report	Purpose
Global AWR	Cluster-wide RAC workload view
Instance AWR	Specific instance performance view

Global AWR is useful for RAC-wide issues. Instance AWR is useful when only one node or instance is slow.

57. What causes one RAC node to be slower than another?

Possible causes:

- Uneven workload distribution
- Service imbalance
- CPU pressure on one node
- Memory pressure
- Local instance wait events

- Interconnect issue on one node
- Storage path issue
- Different optimizer environment
- Different instance parameters
- Blocking sessions on one instance

Check:

```
SELECT inst_id, event, COUNT(*)
FROM gv$session
WHERE status='ACTIVE'
GROUP BY inst_id, event
ORDER BY inst_id, COUNT(*) DESC;
```

58. How do you reduce RAC interconnect traffic?

Options:

- Use service-based workload partitioning.
- Keep related transactions on same instance.
- Avoid cross-instance hot block access.
- Partition hot tables/indexes.
- Use reverse key indexes where suitable.
- Use scalable sequences if available.
- Reduce unnecessary full scans across instances.
- Tune SQL to reduce block visits.
- Separate reporting workload.

59. What is instance affinity?

Instance affinity means directing a workload or service to a preferred RAC instance.

Example:

```
PAYROLL workload runs mainly on Instance 1.
REPORTING workload runs mainly on Instance 2.
```

This reduces block pinging and improves cache locality.

60. What is block pinging?

Block pinging means the same data block is frequently transferred between RAC instances.

This can happen when multiple instances update/read the same hot block.

Symptoms:

- High GC waits
- High interconnect traffic
- Hot segment statistics

- Poor RAC scalability

12. RAC Services, FAN, FCF, TAF, Application Continuity

61. What is FAN?

FAN means Fast Application Notification.

It sends events to clients when RAC services, nodes, or instances go up or down.

Applications can react quickly to failures instead of waiting for TCP timeout.

62. What is FCF?

FCF means Fast Connection Failover.

It uses FAN events to quickly remove dead connections from connection pools and establish new ones.

Commonly used with Oracle connection pools.

63. What is TAF?

TAF means Transparent Application Failover.

It allows client sessions to reconnect to another instance after failure.

TAF can fail over session and SELECT query, but it does not automatically replay DML transactions.

64. What is Application Continuity?

Application Continuity can replay in-flight database requests after recoverable outages.

It improves application availability by hiding some outages from users.

It requires proper services and supported drivers.

65. Difference between TAF and Application Continuity?

TAF	Application Continuity
Older failover feature	Newer request replay feature
Can fail over session/select	Can replay complete request
Limited DML support	Better application transparency
Less intelligent	More advanced

13. RAC Patching and Maintenance

66. How do you patch RAC with minimum downtime?

Use rolling patching where supported.

High-level steps:

- Check patch readme and rolling support.
- Backup Grid and DB homes.
- Patch one node at a time.
- Relocate services from node being patched.
- Stop instance on that node.
- Patch Grid/DB home.
- Start services back.
- Repeat on next node.
- Run datapatch if required.
- Validate cluster and database.

Service relocation:

```
srvctl relocate service -db PRODDB -service APP_SVC -oldinst PRODDB1 -newinst PRODDB2
```

67. What checks do you perform before RAC patching?

Pre-checks:

- Cluster health
- OCR backup
- Voting disk status
- ASM diskgroup status
- Database backup
- Patch conflict check
- OPatch version
- Free space
- Services configuration
- Application failover testing
- Rollback plan

Commands:

```
crsctl check cluster -all
crsctl stat res -t
ocrconfig -manualbackup
opatch lsinventory
```

14. RAC Troubleshooting Scenarios

Scenario 1: One RAC instance is down. What will you do?

Check:

- Cluster resource status.
- Instance alert log.
- CRS logs.
- ASM availability.
- Listener/service status.
- Recent node eviction or crash.
- OS resource issue.
- Try controlled startup if root cause is clear.

Commands:

```
crsctl stat res -t
srvctl status database -db PRODDB
srvctl status instance -db PRODDB -instance PRODDB1
srvctl start instance -db PRODDB -instance PRODDB1
```

Scenario 2: Application cannot connect to RAC. What will you check?

Check in this order:

- SCAN name resolution.
- SCAN listener status.
- Local listener status.
- VIP status.
- Service status.
- Database/instance status.
- Firewall/network.
- TNS connect string.
- Service registration.

Commands:

```
srvctl status scan
srvctl status scan_listener
srvctl status listener
srvctl status service -db PRODDB
lsnrctl status
```

SQL:

```
SELECT inst_id, name FROM gv$services;
```

Scenario 3: SCAN listener is down. What will you do?

Check SCAN status:

```
srvctl status scan
srvctl status scan_listener
```

Start:

```
srvctl start scan
srvctl start scan_listener
```

Check name resolution:

```
nslookup scan_name
```

Check cluster resources:

```
crsctl stat res -t | grep -i scan
```

Scenario 4: VIP is not running on one node. What will you check?

Check:

- Public network interface.
- VIP resource status.
- CRS logs.
- Network configuration.
- DNS/hosts file.
- IP conflict.

Commands:

```
srvctl status vip -node node1
crsctl stat res -t
srvctl start vip -node node1
```

Scenario 5: High GC waits in RAC. How do you resolve?

First identify the type of GC wait.

```
gc current waits = cross-instance modifications
gc cr waits      = cross-instance consistent reads
```

Then check:

- SQL causing waits.
- Hot objects.
- Services placement.
- Instance affinity.
- Interconnect latency.
- LMS CPU usage.
- Segment statistics.

Fix options:

- Move related workload to same service/instance.
- Partition hot objects.
- Use reverse key/scalable sequence for hot index blocks.
- Tune SQL.

- Reduce unnecessary cross-instance access.
- Improve interconnect.

Scenario 6: Node eviction happened. What is your action plan?

Action plan:

- Confirm eviction timestamp.
- Check cluster alert log.
- Check CSSD logs.
- Check OS logs.
- Check CHM/OSWatcher.
- Check private network.
- Check voting disk/storage latency.
- Check CPU/memory starvation.
- Check if eviction was rebootless or full node reboot.
- Prepare RCA.

Do not just restart services without identifying root cause.

Scenario 7: CRS is not starting on one node. What will you check?

Check:

- OHASD status.
- CSSD startup.
- Voting disk access.
- OCR access.
- ASM startup.
- Network configuration.
- File permissions.
- Recent patch/change.

Commands:

```
crsctl check crs
crsctl start crs
ocrcheck
crsctl query css votedisk
```

Logs:

```
$GRID_HOME/log/<hostname>/ohasd/
$GRID_HOME/log/<hostname>/cssd/
$GRID_HOME/log/<hostname>/crsd/
```

Scenario 8: ASM diskgroup is not mounting on one node.

Check:

- ASM instance status.
- Disk visibility from OS.
- ASM disk header status.
- Multipath status.
- ASM alert log.
- Diskgroup compatibility.
- Permissions on devices.

SQL:

```
SELECT name, state FROM v$asm_diskgroup;
SELECT path, header_status, mount_status, mode_status FROM v$asm_disk;
```

Scenario 9: Database service is not failing over.

Check:

- Service preferred/available configuration.
- Service failover type/method.
- Server pool configuration.
- Policy-managed or admin-managed DB.
- FAN/FCF configuration.
- Application connection pool support.
- Service is enabled.

Command:

```
srvctl config service -db PRODDB -service APP_SVC
```

Scenario 10: One node has high CPU.

Check whether workload is balanced.

SQL:

```
SELECT inst_id, COUNT(*)
FROM gv$session
WHERE status='ACTIVE'
GROUP BY inst_id;
```

Top SQL by instance:

```
SELECT inst_id, sql_id, COUNT(*)
FROM gv$active_session_history
WHERE sample_time > SYSDATE - 1/24
GROUP BY inst_id, sql_id
ORDER BY COUNT(*) DESC;
```

Possible causes:

- Services not balanced

- One instance receiving most connections
- Parallel query on one node
- Bad SQL plan on one instance
- OS process consuming CPU
- Application connection pool imbalance

15. Important RAC SQL Queries

Check instance status

```
SELECT inst_id, instance_name, host_name, status
FROM gv$instance;
```

Check database role/open mode

```
SELECT inst_id, name, open_mode, database_role
FROM gv$database;
```

Check services

```
SELECT inst_id, name
FROM gv$services
ORDER BY inst_id, name;
```

Check active sessions by instance

```
SELECT inst_id, status, COUNT(*)
FROM gv$session
GROUP BY inst_id, status
ORDER BY inst_id;
```

Check waits by instance

```
SELECT inst_id, event, COUNT(*)
FROM gv$session
WHERE status='ACTIVE'
GROUP BY inst_id, event
ORDER BY inst_id, COUNT(*) DESC;
```

Check blocking sessions across RAC

```
SELECT inst_id,
       sid,
       serial#,
       blocking_instance,
       blocking_session,
       event,
       sql_id
FROM gv$session
WHERE blocking_session IS NOT NULL;
```

Check global cache waits

```
SELECT inst_id, event, total_waits, time_waited
FROM gv$system_event
```

```
WHERE event LIKE 'gc%'
ORDER BY time_waited DESC;
```

Check hot objects

```
SELECT inst_id,
       owner,
       object_name,
       statistic_name,
       value
FROM gv$segment_statistics
WHERE statistic_name LIKE 'gc%'
AND value > 0
ORDER BY value DESC;
```

16. Common RAC Errors

68. CRS-4535: Cannot communicate with Cluster Ready Services

Meaning: CRSD is not running or not reachable.

Check:

```
crsctl check crs
crsctl stat res -t
```

Logs:

```
$GRID_HOME/log/<hostname>/crsd/
```

69. CRS-4639: Could not contact Oracle High Availability Services

Meaning: OHASD is not running or not reachable.

Action:

```
crsctl check has
crsctl start has
```

70. CRS-4530: Communications failure contacting Cluster Synchronization Services daemon

Meaning: CSSD is down or unavailable.

Check:

```
crsctl check css
```

Also check voting disk, interconnect, and CSSD logs.

71. ORA-29701: Unable to connect to Cluster Synchronization Service

Meaning: Database/ASM cannot connect to CSS.

Possible causes:

- Clusterware down
- CSSD down
- Node not part of cluster
- Grid Infrastructure issue

72. ORA-29740: Evicted by instance number

Meaning: Instance was evicted from RAC cluster.

Check:

- Instance alert log
- CSSD logs
- Cluster alert log
- OS logs
- Interconnect
- CPU/memory/storage issue

17. Level 4 RAC Case Studies

Case Study 1: RAC database is slow only on node 2

I would compare node 2 with other nodes.

Check:

- Active sessions by instance.
- Wait events by instance.
- CPU/memory usage on node 2.
- Top SQL on node 2.
- Services running on node 2.
- Interconnect errors.
- Storage path issue.
- Different execution plans.

SQL:

```
SELECT inst_id, event, COUNT(*)
FROM gv$session
WHERE status='ACTIVE'
GROUP BY inst_id, event
ORDER BY inst_id, COUNT(*) DESC;
```

If only node 2 has high GC waits, check interconnect and hot block access.

If only node 2 has high CPU, check top SQL and service load.

Case Study 2: High `gc current block busy`

This usually means current blocks are hot and being modified across instances.

Check:

- SQL causing wait.
- Object/block involved.
- Segment statistics.
- Application module.
- Insert/update pattern.
- Sequence/index hot block.

Possible fixes:

- Use service affinity.
- Partition workload.
- Partition hot table/index.
- Use reverse key index if right-hand index contention.
- Use scalable sequence.
- Avoid multiple instances updating same hot rows.

Case Study 3: Private interconnect failure

Symptoms:

- GC waits increase
- Node eviction
- CRS/CSSD errors
- Packet loss
- Instance communication failure

Checks:

```
oifcfg getif
crsctl stat res -t
```

Also check OS network bonding, switch logs, NIC errors, MTU mismatch, packet drops, and private network latency.

Case Study 4: Application connections are not balanced

Check:

- Services configuration.
- SCAN listeners.
- REMOTE_LISTENER.
- Connection string uses SCAN.

- Load balancing enabled.
- Application connection pool behavior.
- Runtime Load Balancing/FAN if used.

Command:

```
srvctl config service -db PRODDB -service APP_SVC
```

TNS should use SCAN, not individual node VIPs.

Case Study 5: RAC instance crashed during heavy load

Check:

- Instance alert log.
- Trace files.
- OS logs.
- CRS logs.
- CSSD logs.
- Memory pressure.
- CPU starvation.
- HugePages/Linux memory configuration.
- Storage latency.
- Interconnect errors.

If ORA-29740 appears, investigate eviction.

If ORA-00600/ORA-07445 appears, investigate database bug or corruption.

18. Senior DBA Answer Pattern

For Level 4 RAC interviews, answer like this:

```
1. I first identify whether the issue is database, clusterware, ASM, network, storage, or
application.
2. I check cluster status using crsctl and srvctl.
3. I check GV$ views to compare instances.
4. I check AWR Global and per-instance reports.
5. For RAC waits, I identify SQL, hot objects, and service placement.
6. For eviction, I check CSSD logs, OS logs, CHM/OSWatcher, network, and storage.
7. I fix root cause and validate service/database availability.
```

19. Most Expected Level 4 RAC Questions

1. Explain RAC architecture.
2. What is Clusterware?
3. Difference between OCR and Voting Disk.
4. What is Cache Fusion?
5. What are GCS and GES?

6. What are LMS, LMD, LMON?

7. What is SCAN and why is it used?

8. Difference between VIP and SCAN VIP.

9. What is private interconnect?

10. What causes node eviction?

11. What is split-brain?

12. How do you troubleshoot node eviction?

13. What are common RAC wait events?

14. What causes high `gc current request`?

15. What causes high `gc buffer busy acquire`?

16. How do you reduce interconnect traffic?

17. How do you troubleshoot one slow RAC node?

18. How do you check cluster status?

19. How do you manage RAC services?

20. What is FAN?

21. What is FCF?

22. What is TAF?

23. What is Application Continuity?

24. How do you patch RAC with minimum downtime?

25. What checks are required before RAC patching?

26. How do you troubleshoot SCAN listener issue?

27. How do you troubleshoot ASM diskgroup issue in RAC?

28. How do you identify hot blocks in RAC?

29. What is instance affinity?

30. How do you avoid split-brain?

20. Final Interview Summary

A strong Level 4 RAC answer should sound like this:

```
Oracle RAC provides high availability and scalability by allowing multiple instances to access
the same database on shared storage. Clusterware manages nodes, resources, OCR, voting disks,
VIPs, listeners, services, and failover. RAC uses cache fusion to transfer blocks between
instances through the private interconnect. For performance issues, I compare AWR Global and
instance-level reports, analyze GC waits, identify hot SQL/objects, and verify service
placement. For node eviction, I check CSSD logs, cluster alert log, OS logs, interconnect,
voting disk latency, CPU, memory, and storage. For application connectivity, I verify SCAN,
VIPs, listeners, services, and client connection configuration.
```

That is the expected depth for a Level 4 Oracle RAC DBA interview.